GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:02:18 ; Search time 58.3333 Seconds
(without alignments)
54.421 Million cell updates/sec

Title: US-09-549-186B-7
Perfect score: 104
Sequence: 1 CGSGVRGDFGSLAPRVARQL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	95	91.3	19	24	ABU08642	Foot and mouth dis
2	95	91.3	20	21	AAY94587	Envelope gene epit
3	95	91.3	31	4	AAP30107	Sequence of VP1 ca
4	95	91.3	31	6	AAP50728	Foot-and-mouth dis
5	95	91.3	31	18	AAW17082	Foot and mouth dis
6	95	91.3	35	21	AAY68501	Target antigenic p
7	95	91.3	75	21	AAY68528	Synthetic foot and
8	95	91.3	216	4	AAP30205	Sequence encoded b
9	95	91.3	220	4	AAP30195	Sequence encoded b
10	95	91.3	233	4	AAP30204	Sequence encoded b
11	95	91.3	405	4	AAP30202	Sequence encoded b
12	95	91.3	406	4	AAP30203	Sequence encoded b
13	95	91.3	609	4	AAP30206	Sequence encoded b
14	92	88.5	31	18	AAW17086	Foot and mouth dis
15	91	87.5	31	18	AAW17087	Foot and mouth dis
16	89	85.6	31	18	AAW17084	Foot and mouth dis
17	88	84.6	31	18	AAW17083	Foot and mouth dis
18	88	84.6	31	18	AAW17088	Foot and mouth dis
19	87	83.7	35	21	AAY68502	Target antigenic p
20	87	83.7	46	21	AAY68503	Target antigenic p
21	87	83.7	65	21	AAY68527	Synthetic foot and
22	87	83.7	75	21	AAY68529	Synthetic foot and
23	85	81.7	19	24	ABU08643	Foot and mouth dis
24	85	81.7	35	7	AAP60895	Peptide containing
25	85	81.7	35	9	AAP82746	VP1(A12,119) contg
26	80	76.9	18	4	AAP30110	Sequence of VP1 ca
27	77	74.0	20	14	AAR38543	Foot and Mouth Dis
28	77	74.0	20	14	AAR41793	Foot and Mouth Dis
29	76	73.1	31	18	AAW17085	Foot and mouth dis
30	71	68.3	35	21	AAY68508	Consensus antigeni
31	71	68.3	213	4	AAP30198	Sequence encoded b
32	71	68.3	233	4	AAP30209	Sequence of antige
33	67	64.4	20	5	AAP40103	Sequence at antige
34	67	64.4	216	5	AAP40085	Sequence of foot a
35	64	61.5	248	4	AAP30207	Sequence of an FMD
36	63	60.6	35	21	AAY68509	Consensus antigeni
37	63	60.6	55	21	AAY68533	Synthetic foot and
38	58	55.8	218	4	AAP30196	Sequence encoded b
39	56	53.8	28	4	AAP30316	Sequence of VP1 ca
40	56	53.8	213	21	AAB15428	FMDV antigen for r
41	56	53.8	213	21	AAB15397	Foot-and-mouth dis
42	56	53.8	213	21	AAB15398	Foot-and-mouth dis
43	56	53.8	213	21	AAB15402	Foot-and-mouth dis
44	54	51.9	36	7	AAP60896	Peptide containing
45	54	51.9	36	9	AAP82747	VP1(A24,C) contg.

ALIGNMENTS

RESULT 1
ABU08642
ID ABU08642 standard; peptide; 19 AA.
XX
AC ABU08642;
XX
DT 04-JUN-2003 (first entry)
XX
DE Foot and mouth disease virus (FMDV) VP1-A12141-159, variant FP.
XX
KW Foot and mouth disease virus; FMDV; VP1-A12 region; immunosuppressive;
KW antiinflammatory; antiarthritic; antirheumatic; dermatological;
KW antidiabetic; antianaemic; neuroprotective; vaccine;
KW anti-metallothionein antibody; humoral immune response;
KW autoimmune disease; Grave's disease; rheumatoid arthritis;
KW systemic lupus erythematosus; Type 1 diabetes; pernicious anaemia;
KW multiple sclerosis; Sjorgen's disease.
XX
OS Foot and mouth disease virus.
XX
PN US2003007973-A1.
XX
PD 09-JAN-2003.
XX
PF 24-JUN-2002; 2002US-0178909.
XX
PR 22-JUN-2001; 2001US-300346P.
XX
PA (LYNE/) LYNES M A.
XX
PI Lynes MA;
XX

DR WPI; 2003-353082/33.
XX
PT Treating autoimmune diseases, e.g. Grave's disease, arthritis, lupus,
PT or diabetes by administering a composition consisting of an
PT anti-metallothionein antibody to stimulate a humoral immune response in
PT a subject -
XX
PS Example 7; Fig 9; 24pp; English.
XX
CC The invention describes a method of treating a subject comprising
CC administering to the subject a composition consisting of an
CC anti-metallothionein antibody to stimulate a humoral immune response in
CC a subject. The methods and compositions are useful for treating
CC autoimmune diseases, e.g. Grave's disease, rheumatoid arthritis,
CC systemic lupus erythematosus, Type 1 diabetes, pernicious anaemia,
CC multiple sclerosis or Sjorgen's disease. This is the amino acid
CC sequence of foot and mouth disease virus (FMDV) VP1-A12141-159 region
CC used as an immunogen.
XX
SQ Sequence 19 AA;

Query Match 91.3%; Score 95; DB 24; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GSGVRGDFGSLAPRVARQL 20
||||||||||||||||||||
Db 1 GSGVRGDFGSLAPRVARQL 19

RESULT 2
AAY94587
ID AAY94587 standard; Peptide; 20 AA.
XX
AC AAY94587;
XX
DT 10-JAN-2001 (first entry)
XX
DE Envelope gene epitope of foot and mouth disease virus.
XX
KW Hepatitis B virus nucleocapsid antigen; HBcAg; T cell epitope;
KW cytolytic T lymphocyte; immunogenic; ICE; CTL; FMDV;
KW immunodominant core epitope; immunisation; envelope gene.
XX
OS Foot and mouth disease virus.
XX
PN WO200026385-A1.
XX
PD 11-MAY-2000.
XX
PF 05-NOV-1999; 99WO-US26291.
XX
PR 05-NOV-1998; 98US-0107169.
XX
PA (POWD-) POWDERJECT VACCINES INC.
XX
PI Fuller DL, Fuller JT;
XX
DR WPI; 2000-451623/39.
XX
PT Use of expression vector for nucleic acid immunization that comprises
PT promoter and recombinant nucleic acid sequences encoding Hepatitis B
PT core antigen and T cell epitope from antigen -
XX
PS Example 6; Page 39; 55pp; English.
XX
CC The present invention relates to an immunogeneic recombinant
CC nucleic acid molecule. The molecule consists of a modified hepatitis
CC B virus nucleocapsid antigen (HBcAg) with a T cell epitope sequence
CC inserted within the HBcAg. The creation of a unique restriction site
CC in HBcAg facilitated the insertion of the T cell epitope into the DNA
CC encoding the immunodominant core epitope of the HBcAg. An example of a
CC suitable insertion epitope is the present sequence, the
CC neutralisation epitope from the foot and mouth disease virus envelope
CC gene. Alternatively other T cell epitopes may be inserted
CC (AAY94583, AAY94584, AAY94585, AAY94586, AAY94588). The recombinant
CC nucleic acid molecule may then be used as a reagent in various nucleic
CC acid immunisation strategies. The advantage of this method of
CC immunisation is that the nucleic acid reagents that encode hybrid HBcAg
CC generate an extremely high frequency cellular immune response against
CC the CTL epitope.
XX
SQ Sequence 20 AA;

Query Match 91.3%; Score 95; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GSGVRGDFGSLAPRVARQL 20
||||||||||||||||||||
Db 1 GSGVRGDFGSLAPRVARQL 19

RESULT 3
AAP30107
ID AAP30107 standard; Peptide; 31 AA.
XX
AC AAP30107;
XX
DT 03-APR-1992 (first entry)
XX
DE Sequence of VP1 capisid protein residues 130-160 from the amino-
DE terminus, FMDV, Tubingen type A, subtype 12, strain 119.
XX
KW Antigen; Picornavirus; capsid protein; antibody; detection;
KW vaccine; diagnosis.
XX
OS Foot and mouth disease virus.
XX
PN WO8303547-A.
XX
PD 27-OCT-1983.
XX
PF 14-APR-1983; 83WO-0002644.
XX
PR 25-MAR-1983; 83US-0478847.
PR 14-APR-1982; 82US-0368308.
PR 20-SEP-1984; 84US-0653475.
PR 18-DEC-1984; 84US-0682819.
XX
PA (BITT/) BITTLE J L.
PA (SCRI-) SCRIPPS CLINIC & RE.
XX
PI Bittle JL, Lerner RA;
XX
DR WPI; 1983-807942/44.
XX
PT Antigenic peptide(s) corresp. to picornavirus capsid protein -
PT useful in prodn. of vaccines and in diagnostic tests
XX
PS Example; Page 26; 90pp; English.
XX
CC The peptides of the invention corresp. to a region on the antigenic
CC Picornavirus capsid protein. The capsid protein FMDV VP1 or polio
CC virus VP1. When linked to carriers the peptides are immunogenic.
CC Dose is 20 ug-2mg peptide for inoculations.
XX
SQ Sequence 31 AA;

Query Match 91.3%; Score 95; DB 4; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GSGVRGDFGSLAPRVARQL 20
||||||||||||||||||||
Db 11 GSGVRGDFGSLAPRVARQL 29

```
RESULT 4
AAP50728
ID   AAP50728 standard; protein; 31 AA.
XX
AC   AAP50728;
XX
DT   23-OCT-1991  (first entry)
XX
DE   Foot-and-mouth disease antigenic peptide.
XX
KW   Foot-and-mouth disease; vaccine; antigen;
XX
OS   Foot-and-mouth disease virus.
XX
PN   US4544500-A.
XX
PD   01-OCT-1985.
XX
PF   18-DEC-1984;   84US-0682819.
XX
PR   18-DEC-1984;   84US-0682819.
XX
PA   (SCRI-) SCRIPPS CLINIC & RESEARCH FOUNDATION.
XX
PI   Bittle JL,  Lerner RA;
XX
DR   WPI; 1985-262823/42.
XX
PT   Synthetic foot and mouth disease antigen - comprising 20 amino acid
PT   peptide corresp. to virus protein VP1.
XX
PS   Disclosure; page 4; 7pp; English.
XX
CC   The peptide corresponds to amino acids 130-160 of the foot and mouth
CC   disease virus VP1 protein. It represents a monospecific synthetic
CC   antigenic determinant. The peptide can be used to vaccinate animals,
CC   esp. swine and cattle, against the virus with a single inoculation.
XX
SQ   Sequence   31 AA;

  Query Match             91.3%;  Score 95;  DB 6;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-07;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||||||||||||||||||
Db         11 GSGVRGDFGSLAPRVARQL 29


RESULT 5
AAW17082
ID   AAW17082 standard; peptide; 31 AA.
XX
AC   AAW17082;
XX
DT   13-JUN-1997  (first entry)
XX
DE   Foot and mouth disease virus G-H loop of protein VP-1.
XX
KW   Foot and mouth disease; aphthovirus; vaccine; antigen; epitope;
KW   immunogen; G-H loop; VP-1; deletion; mutant; binding site removal;
KW   antibody; non-infectious; attenuated.
XX
OS   Foot and mouth disease virus.
XX
PN   US5612040-A.
XX
PD   18-MAR-1997.
XX
PF   07-APR-1995;   95US-0418716.
XX
PR   07-APR-1995;   95US-0418716.
XX
PA   (USDA ) US SEC OF AGRIC.
XX
PI   Baxt B,  Berinstein A,  Kang AS,  Mason PW,  Reider E;
XX
DR   WPI; 1997-192081/17.
XX
PT   Genetically modified foot-and-mouth disease virus - lacks
PT   cell-binding site and is non-infectious, useful in vaccines
XX
PS   Disclosure; Figure 1; 10pp; English.
XX
CC   AAW17082 represents amino acids 130-163 of wild-type foot and mouth
CC   disease (FMD) virus  VP-1 structural protein. The sequence represents
CC   the G-H loop, a flexible loop between the G and H beta strands of VP-1.
CC   The G-H loop is the main antigenic site of VP-1 and is also involved in
CC   cell binding. The FMD virus G-H loop peptide can be modified in a FMD
CC   virus by deletion of the amino acid (aa) sequence
CC   Gly-Val-Arg-Gly-Asp-Phe (aa 142-147) and replacing this sequence with the
CC   sequence Asn-Pro. This results in the virus retaining its immunogenicity
CC   but is not infectious because the cell binding site has been removed. The
CC   genetically modified FMD virus is useful in vaccines against the
CC   disease.
XX
SQ   Sequence   31 AA;

  Query Match             91.3%;  Score 95;  DB 18;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-07;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||||||||||||||||||
Db         11 GSGVRGDFGSLAPRVARQL 29


RESULT 6
AAY68501
ID   AAY68501 standard; peptide; 35 AA.
XX
AC   AAY68501;
XX
DT   05-MAY-2000  (first entry)
XX
DE   Target antigenic peptide derived from amino acids 134-169 of VP1.
XX
KW   Antigenic peptide; VP1 capsid protein; FMDV strain A12; epitope;
KW   helper T-cell epitope; immune response; vaccine; FMDV infection.
XX
OS   Foot and mouth disease virus.
XX
PN   WO9966954-A1.
XX
PD   29-DEC-1999.
XX
PF   21-JUN-1999;   99WO-US13921.
XX
PR   20-JUN-1998;   98US-0100600.
XX
PA   (UNBI-) UNITED BIOMEDICAL INC.
XX
PI   Wang CY,  Shen M;
XX
DR   WPI; 2000-160563/14.
XX
PT   Synthetic peptide used in protecting animals against Foot-and-Mouth
PT   Disease Virus infections comprises a  helper T-cell epitope and a
PT   Foot-and-Mouth Disease Virus epitope conjugate -
XX
PS   Claim 1; Page 45; 115pp; English.
XX
CC   The present sequence represents a target antigenic peptide derived from
CC   the VP1 capsid protein of Foot and Mouth Disease Virus (FMDV) strain A12.
```

```
CC   The peptide is used to design synthetic epitopes, which are conjugated
CC   to a helper T-cell epitope and optionally an immunostimulator sequence.
CC   The peptides are covalently linked to each other, preferably by peptide
CC   bonds generated through direct synthesis. The synthetic vaccines of the
CC   invention require lower levels of biosecurity, have a reduced loss of
CC   immunogenic activity, are less prone to product instability and
CC   lot-to-lot variability and do not need periodic revision due to
CC   antigenic variation in the field. The peptides are used as immunogens
CC   in vaccines to generate an enhanced immune response against a FMDV
CC   antigen in animals, such as swine, cattle, sheep, goats and susceptible
CC   wild species. They can also be used to diagnose FMDV infection in a
CC   mammal.
XX
SQ   Sequence   35 AA;

  Query Match             91.3%;  Score 95;  DB 21;  Length 35;
  Best Local Similarity   100.0%;  Pred. No. 2e-07;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||||||||||||||||
Db          7 GSGVRGDFGSLAPRVARQL 25


RESULT 7
AAY68528
ID   AAY68528 standard; peptide; 75 AA.
XX
AC   AAY68528;
XX
DT   05-MAY-2000  (first entry)
XX
DE   Synthetic foot and mouth disease virus immunogen.
XX
KW   Antigenic peptide; VP1 capsid protein; serotype Asia; epitope;
KW   helper T-cell epitope; immune response; vaccine; FMDV infection.
XX
OS   Synthetic.
OS   Yersinia sp.
OS   Foot and mouth disease virus.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..16
FT                   /note= "Yersinia invasin domain (immunostimulator
FT                    peptide"
FT   Peptide         17..18
FT                   /note= "spacer"
FT   Peptide         19..38
FT                   /note= "autologous helper T-cell epitope derived
FT                    from VP1 capsid protein"
FT   Peptide         39..40
FT                   /note= "spacer"
FT   Peptide         41..75
FT                   /note= "synthetic VP1 capsid protein epitope"
XX
PN   WO9966954-A1.
XX
PD   29-DEC-1999.
XX
PF   21-JUN-1999;   99WO-US13921.
XX
PR   20-JUN-1998;   98US-0100600.
XX
PA   (UNBI-) UNITED BIOMEDICAL INC.
XX
PI   Wang CY,  Shen M;
XX
DR   WPI; 2000-160563/14.
XX
PT   Synthetic peptide used in protecting animals against Foot-and-Mouth
PT   Disease Virus infections comprises a  helper T-cell epitope and a
PT   Foot-and-Mouth Disease Virus epitope conjugate -
XX
PS   Claim 8; Page 96; 115pp; English.
XX
CC   The present sequence represents a foot and mouth disease (FMDV) vaccine
CC   of the invention. The specification describes epitopes derived from an
CC   antigen of the VP1 capsid protein of FMDV. The VP1 epitopes are
CC   conjugated to a helper T-cell epitope and optionally an immunostimulator
CC   sequence. The peptides are covalently linked to each other, preferably
CC   by peptide bonds generated through direct synthesis. The synthetic
CC   vaccines of the invention require lower levels of biosecurity, have a
CC   reduced loss of immunogenic activity, are less prone to product
CC   instability and lot-to-lot variability and do not need periodic revision
CC   due to antigenic variation in the field. The peptides are used as
CC   immunogens in vaccines to generate an enhanced immune response against
CC   a FMDV antigen in animals, such as swine, cattle, sheep, goats and
CC   susceptible wild species. They can also be used to diagnose FMDV
CC   infection in a mammal.
XX
SQ   Sequence   75 AA;

  Query Match             91.3%;  Score 95;  DB 21;  Length 75;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-07;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||||||||||||||||
Db         47 GSGVRGDFGSLAPRVARQL 65


RESULT 8
AAP30205
ID   AAP30205 standard; Protein; 216 AA.
XX
AC   AAP30205;
XX
DT   29-JUL-1992  (first entry)
XX
DE   Sequence encoded by direct expression vector pFM 10 which
DE   comprises the promoter and operator of the E. coli tryptophan
DE   operon AAs 1-211 of the VP3 gene linked to 4 AAs from pBR322.
XX
KW   Vaccine; immunogen; antigen; viral protein.
XX
OS   Foot and mouth disease virus.
XX
PN   EP68693-A.
XX
PD   05-JAN-1983.
XX
PF   11-JUN-1982;   82EP-0303040.
XX
PR   04-MAY-1982;   82US-0374855.
PR   16-JUN-1981;   81US-0274103.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Kleid DG,  Yansura DG;
XX
DR   WPI; 1983-05055K/03.
DR   N-PSDB; AAN30144.
XX
PT   Polypeptide antigens of foot-and-mouth disease - obtd. by
PT   recombinant DNA technology for vaccine prodn.
XX
PS   Example; Page 36-37; 81pp; English.
XX
CC   The inventors claim a polypeptide comprising at least one antigen of
CC   FMDV, or a fusion polypeptide of at least two antigens from at least
CC   two different FMDV strands, and encoding DNA.  Pref. the fusion
CC   polypeptide has the sequence of two of the A, C, O, Asia-1, SAT 1,
CC   SAT 2 and SAT 3 type strains of the virus.  The polypeptides pref.
CC   comprise AAs 55-177 or 1-210 of the VP3 virus.  Fusion polypeptides
CC   may comprise AAs 1-210 fused to 8-210 of VP3 virus protein.  pFM1,
CC   pFM2, pFM3, pFM10, pFM20, pFMB1, pFMC, pFMD, pFMF and pFMG are also
```

```
CC   claimed.
XX
SQ   Sequence   216 AA;

  Query Match             91.3%;  Score 95;  DB 4;  Length 216;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-06;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||||||||||||||||||
Db        141 GSGVRGDFGSLAPRVARQL 159


RESULT 9
AAP30195
ID   AAP30195 standard; Protein; 220 AA.
XX
AC   AAP30195;
XX
DT   29-JUL-1992  (first entry)
XX
DE   Sequence encoded by VP3 gene of FMDV type A12 (T465).
XX
KW   Vaccine; immunogen; antigen; viral protein.
XX
OS   Foot and mouth disease virus.
XX
PN   EP68693-A.
XX
PD   05-JAN-1983.
XX
PF   11-JUN-1982;   82EP-0303040.
XX
PR   04-MAY-1982;   82US-0374855.
PR   16-JUN-1981;   81US-0274103.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Kleid DG,  Yansura DG;
XX
DR   WPI; 1983-05055K/03.
DR   N-PSDB; AAN30127.
XX
PT   Polypeptide antigens of foot-and-mouth disease - obtd. by
PT   recombinant DNA technology for vaccine prodn.
XX
PS   Disclosure; Fig 2; 81pp; English.
XX
CC   The inventors claim a polypeptide comprising at least one antigen of
CC   FMDV, or a fusion polypeptide of at least two antigens from at least
CC   two different FMDV strands, and encoding DNA.  Pref. the fusion
CC   polypeptide has the sequence of two of the A, C, O, Asia-1, SAT 1,
CC   SAT 2 and SAT 3 type strains of the virus.  The polypeptides pref.
CC   comprise AAs 55-177 or 1-210 of the VP3 virus.  Fusion polypeptides
CC   may comprise AAs 1-210 fused to 8-210 of VP3 virus protein.  pFM1,
CC   pFM2, pFM3, pFM10, pFM20, pFMB1, pFMC, pFMD, pFMF and pFMG are also
CC   claimed.
XX
SQ   Sequence   220 AA;

  Query Match             91.3%;  Score 95;  DB 4;  Length 220;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-06;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||||||||||||||||||
Db        140 GSGVRGDFGSLAPRVARQL 158


RESULT 10
AAP30204
ID   AAP30204 standard; Protein; 233 AA.
XX
AC   AAP30204;
XX
DT   29-JUL-1992  (first entry)
XX
DE   Sequence encoded by pFM3 which comprises 17 AAs from the LE' gene
DE   construction (fragment 3a), linked to a methionine AA (CNBr cleavable),
DE   linked to the AAs 1-211 of the VP3 FMDV A12, linked to 4 AAs from
DE   pBR322.
XX
KW   Vaccine; immunogen; antigen; viral protein.
XX
OS   Foot and mouth disease virus and Escherichia coli.
XX
PN   EP68693-A.
XX
PD   05-JAN-1983.
XX
PF   11-JUN-1982;   82EP-0303040.
XX
PR   04-MAY-1982;   82US-0374855.
PR   16-JUN-1981;   81US-0274103.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Kleid DG,  Yansura DG;
XX
DR   WPI; 1983-05055K/03.
DR   N-PSDB; AAN30142.
XX
PT   Polypeptide antigens of foot-and-mouth disease - obtd. by
PT   recombinant DNA technology for vaccine prodn.
XX
PS   Example; Page 33-34; 81pp; English.
XX
CC   The inventors claim a polypeptide comprising at least one antigen of
CC   FMDV, or a fusion polypeptide of at least two antigens from at least
CC   two different FMDV strands, and encoding DNA.  Pref. the fusion
CC   polypeptide has the sequence of two of the A, C, O, Asia-1, SAT 1,
CC   SAT 2 and SAT 3 type strains of the virus.  The polypeptides pref.
CC   comprise AAs 55-177 or 1-210 of the VP3 virus.  Fusion polypeptides
CC   may comprise AAs 1-210 fused to 8-210 of VP3 virus protein.  pFM1,
CC   pFM2, pFM3, pFM10, pFM20, pFMB1, pFMC, pFMD, pFMF and pFMG are also
CC   claimed.
XX
SQ   Sequence   233 AA;

  Query Match             91.3%;  Score 95;  DB 4;  Length 233;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-06;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||||||||||||||||||
Db        158 GSGVRGDFGSLAPRVARQL 176


RESULT 11
AAP30202
ID   AAP30202 standard; Protein; 405 AA.
XX
AC   AAP30202;
XX
DT   29-JUL-1992  (first entry)
XX
DE   Sequence encoded by pFM1 which comprises 190 AAs form the LE' gene
DE   construction linked to 6 AAs from the linker molecule (fragment 4),
DE   linked to 205 AAs from the VP3 FMDV A12 (codons 7-211), linked to
DE   4 AAs from pBR322.
XX
KW   Vaccine; immunogen; antigen; viral protein.
XX
OS   Foot and mouth disease virus and Escherichia coli.
XX
PN   EP68693-A.
```

XX
PD 05-JAN-1983.
XX
PF 11-JUN-1982; 82EP-0303040.
XX
PR 04-MAY-1982; 82US-0374855.
PR 16-JUN-1981; 81US-0274103.
XX
PA (GETH) GENENTECH INC.
XX
PI Kleid DG, Yansura DG;
XX
DR WPI; 1983-05055K/03.
DR N-PSDB; AAN30140.
XX
PT Polypeptide antigens of foot-and-mouth disease - obtd. by
PT recombinant DNA technology for vaccine prodn.
XX
PS Example; Page 29-31; 81pp; English.
XX
CC The inventors claim a polypeptide comprising at least one antigen of
CC FMDV, or a fusion polypeptide of at least two antigens from at least
CC two different FMDV strands, and encoding DNA. Pref. the fusion
CC polypeptide has the sequence of two of the A, C, O, Asia-1, SAT 1,
CC SAT 2 and SAT 3 type strains of the virus. The polypeptides pref.
CC comprise AAs 55-177 or 1-210 of the VP3 virus. Fusion polypeptides
CC may comprise AAs 1-210 fused to 8-210 of VP3 virus protein. pFM1,
CC pFM2, pFM3, pFM10, pFM20, pFMB1, pFMC, pFMD, pFMF and pFMG are also
CC claimed.
XX
SQ Sequence 405 AA;

Query Match 91.3%; Score 95; DB 4; Length 405;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GSGVRGDFGSLAPRVARQL 20
||||||||||||||||||||
Db 330 GSGVRGDFGSLAPRVARQL 348

RESULT 12
AAP30203
ID AAP30203 standard; Protein; 406 AA.
XX
AC AAP30203;
XX
DT 29-JUL-1992 (first entry)
XX
DE Sequence encoded by pFM2 which comprises190 AAs from the LE' gene,
DE linked to a methionine AA (CNBr cleavable), linked to the AAs 1-211
DE of the VP3 FMDV A12 (codons 1-7 are derived form synthetic DNA, 8-211
DE are natural codons), linked to 4 amino acids from pBR322.
XX
KW Vaccine; immunogen; antigen; viral protein.
XX
OS Foot and mouth disease virus and Escherichia coli.
XX
PN EP68693-A.
XX
PD 05-JAN-1983.
XX
PF 11-JUN-1982; 82EP-0303040.
XX
PR 04-MAY-1982; 82US-0374855.
PR 16-JUN-1981; 81US-0274103.
XX
PA (GETH) GENENTECH INC.
XX
PI Kleid DG, Yansura DG;
XX
DR WPI; 1983-05055K/03.
DR N-PSDB; AAN30141.
XX
PT Polypeptide antigens of foot-and-mouth disease - obtd. by
PT recombinant DNA technology for vaccine prodn.
XX
PS Example; Page 31-32; 81pp; English.
XX
CC The inventors claim a polypeptide comprising at least one antigen of
CC FMDV, or a fusion polypeptide of at least two antigens from at least
CC two different FMDV strands, and encoding DNA. Pref. the fusion
CC polypeptide has the sequence of two of the A, C, O, Asia-1, SAT 1,
CC SAT 2 and SAT 3 type strains of the virus. The polypeptides pref.
CC comprise AAs 55-177 or 1-210 of the VP3 virus. Fusion polypeptides
CC may comprise AAs 1-210 fused to 8-210 of VP3 virus protein. pFM1,
CC pFM2, pFM3, pFM10, pFM20, pFMB1, pFMC, pFMD, pFMF and pFMG are also
CC claimed.
XX
SQ Sequence 406 AA;

Query Match 91.3%; Score 95; DB 4; Length 406;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GSGVRGDFGSLAPRVARQL 20
||||||||||||||||||||
Db 331 GSGVRGDFGSLAPRVARQL 349

RESULT 13
AAP30206
ID AAP30206 standard; Protein; 609 AA.
XX
AC AAP30206;
XX
DT 29-JUL-1992 (first entry)
XX
DE Sequence encoded by pFM 20 which is the polyantigen comprising 190 AAs
DE coded be the LE' gene construction, linked to a methionine AA (CNBr
DE cleavable), linked to the AAs 1-211 of the VP3 FMDV A12 (codons 1-7
DE are derived from synthetic DNA), linked to AAs 8-211 of the VP3
DE FMDV A12, linked to 4 AAs from pBR322.
XX
KW Vaccine; immunogen; antigen; viral protein.
XX
OS Foot and mouth disease virus and Escherichia coli.
XX
PN EP68693-A.
XX
PD 05-JAN-1983.
XX
PF 11-JUN-1982; 82EP-0303040.
XX
PR 04-MAY-1982; 82US-0374855.
PR 16-JUN-1981; 81US-0274103.
XX
PA (GETH) GENENTECH INC.
XX
PI Kleid DG, Yansura DG;
XX
DR WPI; 1983-05055K/03.
DR N-PSDB; AAN30145.
XX
PT Polypeptide antigens of foot-and-mouth disease - obtd. by
PT recombinant DNA technology for vaccine prodn.
XX
PS Example; Page 38-41; 81pp; English.
XX
CC The inventors claim a polypeptide comprising at least one antigen of
CC FMDV, or a fusion polypeptide of at least two antigens from at least
CC two different FMDV strands, and encoding DNA. Pref. the fusion
CC polypeptide has the sequence of two of the A, C, O, Asia-1, SAT 1,
CC SAT 2 and SAT 3 type strains of the virus. The polypeptides pref.
CC comprise AAs 55-177 or 1-210 of the VP3 virus. Fusion polypeptides
CC may comprise AAs 1-210 fused to 8-210 of VP3 virus protein. pFM1,

```
CC   pFM2, pFM3, pFM10, pFM20, pFMB1, pFMC, pFMD, pFMF and pFMG are also
CC   claimed.
XX
SQ   Sequence   609 AA;

  Query Match             91.3%;  Score 95;  DB 4;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-06;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||||||||||||||||||
Db        331 GSGVRGDFGSLAPRVARQL 349


RESULT 14
AAW17086
ID   AAW17086 standard; peptide; 31 AA.
XX
AC   AAW17086;
XX
DT   13-JUN-1997  (first entry)
XX
DE   Foot and mouth disease virus modified G-H loop of protein VP-1.
XX
KW   Foot and mouth disease; aphthovirus; vaccine; antigen; epitope;
KW   immunogen; G-H loop; VP-1; deletion; mutant; binding site removal;
KW   antibody; non-infectious; attenuated.
XX
OS   Foot and mouth disease virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 15
FT                   /label= substitution
FT                   /note= "wild-type Arg replaced with Lys"
XX
PN   US5612040-A.
XX
PD   18-MAR-1997.
XX
PF   07-APR-1995;   95US-0418716.
XX
PR   07-APR-1995;   95US-0418716.
XX
PA   (USDA ) US SEC OF AGRIC.
XX
PI   Baxt B,  Berinstein A,  Kang AS,  Mason PW,  Reider E;
XX
DR   WPI; 1997-192081/17.
XX
PT   Genetically modified foot-and-mouth disease virus - lacks
PT   cell-binding site and is non-infectious, useful in vaccines
XX
PS   Disclosure; Figure 2; 10pp; English.
XX
CC   AAW17083-W17088 are peptides representing mutant versions of amino acids
CC   130-163 of wild-type foot and mouth disease (FMD) virus VP-1 structural
CC   protein. The mutants were made in order to investigate the requirements
CC   of this region for cell binding. The sequences are mutant versions of
CC   the G-H loop, a flexible loop between the G and H beta strands of VP-1)
CC   VP-1, which is the main antigenic site of VP-1 and is also involved in
CC   cell binding. The FMD virus G-H loop peptide can be modified in a FMD
CC   virus by deletion of the amino acid (aa) sequence
CC   Gly-Val-Arg-Gly-Asp-Phe (aa 142-147) and replacing this sequence with the
CC   sequence Asn-Pro. This results in the virus retaining its immunogenicity
CC   but is not infectious because the cell binding site has been removed. The
CC   genetically modified FMD virus is useful in vaccines against the
CC   disease.
XX
SQ   Sequence   31 AA;

  Query Match             88.5%;  Score 92;  DB 18;  Length 31;
  Best Local Similarity   94.7%;  Pred. No. 5.2e-07;
  Matches   18;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||:|||||||||||||||
Db         11 GSGVKGDFGSLAPRVARQL 29


RESULT 15
AAW17087
ID   AAW17087 standard; peptide; 31 AA.
XX
AC   AAW17087;
XX
DT   13-JUN-1997  (first entry)
XX
DE   Foot and mouth disease virus modified G-H loop of protein VP-1.
XX
KW   Foot and mouth disease; aphthovirus; vaccine; antigen; epitope;
KW   immunogen; G-H loop; VP-1; deletion; mutant; binding site removal;
KW   antibody; non-infectious; attenuated.
XX
OS   Foot and mouth disease virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 17
FT                   /label= substitution
FT                   /note= "wild-type Asp replaced with Glu"
XX
PN   US5612040-A.
XX
PD   18-MAR-1997.
XX
PF   07-APR-1995;   95US-0418716.
XX
PR   07-APR-1995;   95US-0418716.
XX
PA   (USDA ) US SEC OF AGRIC.
XX
PI   Baxt B,  Berinstein A,  Kang AS,  Mason PW,  Reider E;
XX
DR   WPI; 1997-192081/17.
XX
PT   Genetically modified foot-and-mouth disease virus - lacks
PT   cell-binding site and is non-infectious, useful in vaccines
XX
PS   Disclosure; Figure 2; 10pp; English.
XX
CC   AAW17083-W17088 are peptides representing mutant versions of amino acids
CC   130-163 of wild-type foot and mouth disease (FMD) virus VP-1 structural
CC   protein. The mutants were made in order to investigate the requirements
CC   of this region for cell binding. The sequences are mutant versions of
CC   the G-H loop, a flexible loop between the G and H beta strands of VP-1)
CC   VP-1, which is the main antigenic site of VP-1 and is also involved in
CC   cell binding. The FMD virus G-H loop peptide can be modified in a FMD
CC   virus by deletion of the amino acid (aa) sequence
CC   Gly-Val-Arg-Gly-Asp-Phe (aa 142-147) and replacing this sequence with the
CC   sequence Asn-Pro. This results in the virus retaining its immunogenicity
CC   but is not infectious because the cell binding site has been removed. The
CC   genetically modified FMD virus is useful in vaccines against the
CC   disease.
XX
SQ   Sequence   31 AA;

  Query Match             87.5%;  Score 91;  DB 18;  Length 31;
  Best Local Similarity   94.7%;  Pred. No. 7.5e-07;
  Matches   18;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||||:|||||||||||||
Db         11 GSGVRGEFGSLAPRVARQL 29


Search completed: October  9, 2003, 10:14:18
Job time : 60.3333 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:10:09 ; Search time 19.6667 Seconds
(without alignments)
97.799 Million cell updates/sec

Title: US-09-549-186B-7
Perfect score: 104
Sequence: 1 CGSGVRGDFGSLAPRVARQL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	95	91.3	2332	1	GNNY4F	genome polyprotein
2	67	64.4	216	2	A03911	genome polyprotein
3	50	48.1	230	2	A03909	genome polyprotein
4	50	48.1	2336	2	S37077	genome polyprotein
5	48	46.2	377	2	A53384	polysialic acid ca
6	48	46.2	377	2	S60758	siaA protein - Nei
7	48	46.2	2333	1	GNNY2F	genome polyprotein
8	47	45.2	351	2	E75631	iron ABC transport
9	47	45.2	365	2	F96017	hypothetical prote
10	46.5	44.7	326	2	D70696	probable dtdp-gluc
11	46	44.2	109	2	S61252	genome polyprotein
12	46	44.2	109	2	S61253	genome polyprotein
13	46	44.2	267	2	S30395	ribosomal protein
14	46	44.2	564	2	F90965	hypothetical prote
15	46	44.2	564	2	F85813	hypothetical prote
16	46	44.2	569	2	H64959	probable membrane
17	46	44.2	573	2	C71264	hypothetical prote
18	46	44.2	651	2	T32875	hypothetical prote
19	45.5	43.8	429	2	S45459	TOM34 protein - ye
20	45	43.3	303	2	E72463	hypothetical prote
21	44.5	42.8	1080	2	A35088	phycobilisome link
22	44	42.3	76	2	D82844	carbon storage reg
23	44	42.3	150	2	S60258	meltrin beta - mou
24	44	42.3	379	1	D69080	Cdc6 related prote
25	43.5	41.8	368	2	F72509	probable developme
26	43	41.3	301	2	T16844	hypothetical prote
27	43	41.3	312	2	A56911	TRADD protein - hu
28	42	40.4	58	2	S66328	protein kinase AK1
29	42	40.4	291	2	T37992	probable tricarbox
30	42	40.4	296	2	H75557	hypothetical prote
31	42	40.4	308	2	AC3604	n-acetylglucosamin
32	42	40.4	309	2	H71328	probable flagellar
33	42	40.4	331	2	E82389	probable outer mem
34	42	40.4	406	2	S40774	ribonucleoprotein
35	42	40.4	467	2	B75322	probable oligoendo
36	42	40.4	486	1	KRXL	keratin 3, type I,
37	42	40.4	529	2	T23740	hypothetical prote
38	42	40.4	630	2	JC7831	cancer-associated
39	42	40.4	795	2	E86247	receptor-like prot
40	42	40.4	809	1	S60112	transitional endop
41	42	40.4	810	2	S75931	hypothetical prote
42	42	40.4	978	2	E96787	protein T4O12.5 [i
43	42	40.4	990	2	T03784	probable receptor
44	42	40.4	992	2	T05335	hypothetical prote
45	42	40.4	1002	2	T46033	receptor protein k

ALIGNMENTS

RESULT 1
GNNY4F
genome polyprotein - foot-and-mouth disease virus A (strain A12)
N;Contains: coat protein VP1; coat protein VP2; coat protein VP3; coat protein VP4; c
tein VPg2; genome-linked protein VPg3; nonstructural protein p20a; proteinase (EC 3.4
C;Species: Aphthovirus A (foot-and-mouth disease virus A)
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C;Accession: A25794
R;Robertson, B.H.; Grubman, M.J.; Weddell, G.N.; Moore, D.M.; Welsh, J.D.; Fischer, T
J. Virol. 54, 651-660, 1985
A;Title: Nucleotide and amino acid sequence coding for polypeptides of foot-and-mouth
A;Reference number: A25794; MUID:85211015; PMID:2987518
A;Accession: A25794
A;Molecule type: genomic RNA
A;Residues: 1-2332 <ROB>
A;Cross-references: GB:M10975; NID:g210306; PIDN:AAA42593.1; PID:g210307
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: coat protein; core protein; genome-linked protein; hydrolase; nonstructur
F;1-216/Product: nonstructural protein p20a #status predicted <NPA>
F;217-285/Product: coat protein VP4 #status predicted <VP4>
F;286-503/Product: coat protein VP2 #status predicted <VP2>
F;504-723/Product: coat protein VP3 #status predicted <VP3>
F;724-937/Product: coat protein VP1 #status predicted <VP1>
F;938-953/Product: core protein X #status predicted <CPX>
F;954-1107/Product: core protein p14 #status predicted <C14>
F;1108-1425/Product: core protein p41 #status predicted <C41>
F;1426-1578/Product: core protein p19 #status predicted <C19>
F;1579-1601/Product: genome-linked protein VPg1 #status predicted <VG1>
F;1602-1625/Product: genome-linked protein VPg2 #status predicted <VG2>
F;1626-1649/Product: genome-linked protein VPg3 #status predicted <VG3>
F;1650-1862/Product: proteinase #status predicted <PTS>
F;1863-2332/Product: RNA-directed RNA polymerase #status predicted <RRP>

Query Match 91.3%; Score 95; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||||||||||||||||||
Db        864 GSGVRGDFGSLAPRVARQL 882
```

RESULT 2
A03911
genome polyprotein - foot-and-mouth disease virus A (strain A24 Cruzeiro) (fragment)
N;Contains: coat protein VP1; core protein p52
C;Species: Aphthovirus A (foot-and-mouth disease virus A)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 26-Aug-1999
C;Accession: A03911
R;Makoff, A.J.; Paynter, C.A.; Rowlands, D.J.; Boothroyd, J.C.
Nucleic Acids Res. 10, 8285-8295, 1982
A;Title: Comparison of the amino acid sequence of the major immunogen from three sero

A;Reference number: A37503; MUID:83143292; PMID:6298715
A;Accession: A03911
A;Molecule type: genomic RNA
A;Residues: 1-216 <MAK>
A;Cross-references: GB:J02183; NID:g210312; PIDN:AAA42596.1; PID:g210313
C;Comment: Specific enzymatic cleavage in vivo yields mature proteins including coat pro
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: polyprotein

Query Match 64.4%; Score 67; DB 2; Length 216;
Best Local Similarity 73.7%; Pred. No. 0.0051;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GSGVRGDFGSLAPRVARQL 20
||| ||| |||| || :||
Db 129 GSGRRGDMGSLAARVVKQL 147

RESULT 3
A03909
genome polyprotein - foot-and-mouth disease virus A (strain A5) (fragment)
N;Contains: coat protein VP1; coat protein VP3; core protein p52
C;Species: Aphthovirus A (foot-and-mouth disease virus A)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 07-Feb-1997
C;Accession: A03909
R;Beck, E.; Feil, G.; Strohmaier, K.
EMBO J. 2, 555-559, 1983
A;Title: The molecular basis of the antigenic variation of foot-and-mouth disease virus.
A;Reference number: A03909; MUID:84028562; PMID:6194987
A;Accession: A03909
A;Molecule type: mRNA
A;Residues: 1-230 <BEC>
A;Note: the authors translated the codon GCA for residue 114 as Val
C;Comment: Specific enzymatic cleavages in vivo yield mature proteins including the carb
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: polyprotein

Query Match 48.1%; Score 50; DB 2; Length 230;
Best Local Similarity 57.9%; Pred. No. 2.6;
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 2 GSGVRGDFGSLAPRVARQL 20
| ||| || | | |:||
Db 143 GGPRRGDMGSAAARAAKQL 161

RESULT 4
S37077
genome polyprotein - foot-and-mouth disease virus A (strain A22/550 Azerbaijan 65)
N;Contains: coat protein VP1; coat protein VP2; coat protein VP3; coat protein VP4; core
tein VPg2; genome-linked protein VPg3; nonstructural protein p20a; proteinase (EC 3.4.-.
C;Species: Aphthovirus A (foot-and-mouth disease virus A)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 26-Aug-1999
C;Accession: S37077; JN0413
R;Sosnovtsev, S.V.; Onischenko, A.M.; Petrov, N.A.; Kalashnikova, T.I.; Mamaeva, N.V.; D
submitted to the EMBL Data Library, August 1993
A;Reference number: S37077
A;Accession: S37077
A;Molecule type: genomic RNA
A;Residues: 1-2336 <SOS>
A;Cross-references: EMBL:X74812; NID:g397965; PIDN:CAA52812.1; PID:g397966
R;Onishchenko, A.M.; Petrov, N.A.; Blinov, V.M.; Vassilenko, S.K.; Sandakhchiev, L.S.; B
Bioorg. Khim. 12, 416-419, 1986
A;Title: The DNA-copy primary structure for the gene of the FMDV A22 serotype VP1-protei
A;Reference number: JN0413; MUID:86186952; PMID:2421736
A;Accession: JN0413
A;Molecule type: genomic RNA
A;Residues: 702-955 <ONI>
A;Cross-references: GB:M38362; NID:g210514; PIDN:AAA42664.1; PID:g210515
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: coat protein; core protein; genome-linked protein; hydrolase; nonstructural
F;1-217/Product: nonstructural protein p20a #status predicted <NPA>
F;218-286/Product: coat protein VP4 #status predicted <VP4>
F;287-504/Product: coat protein VP2 #status predicted <VP2>
F;505-724/Product: coat protein VP3 #status predicted <VP3>
F;725-938/Product: coat protein VP1 #status predicted <VP1>
F;939-954/Product: core protein X #status predicted <CPX>
F;955-1108/Product: core protein p14 #status predicted <C14>
F;1109-1426/Product: core protein p41 #status predicted <C41>
F;1427-1579/Product: core protein p19 #status predicted <C19>
F;1580-1602/Product: genome-linked protein VPg1 #status predicted <VG1>
F;1603-1626/Product: genome-linked protein VPg2 #status predicted <VG2>
F;1627-1650/Product: genome-linked protein VPg3 #status predicted <VG3>
F;1651-1863/Product: proteinase #status predicted <PTS>
F;1864-2333/Product: RNA-directed RNA polymerase #status predicted <RRP>

Query Match 48.1%; Score 50; DB 2; Length 2336;
Best Local Similarity 63.2%; Pred. No. 26;
Matches 12; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 GSGVRGDFGSLAPRVARQL 20
| | ||| || ||| ||
Db 865 GMGRRGDLEPLAARVAAQL 883

RESULT 5
A53384
polysialic acid capsule biosynthesis protein SynX NMB0070 [imported] - Neisseria meni
C;Species: Neisseria meningitidis
C;Date: 25-Aug-1995 #sequence_revision 25-Aug-1995 #text_change 19-Jan-2001
C;Accession: A53384; D81241
R;Swartley, J.S.; Stephens, D.S.
J. Bacteriol. 176, 1530-1534, 1994
A;Title: Identification of a genetic locus involved in the biosynthesis of N-acetyl-D
.
A;Reference number: A53384; MUID:94156865; PMID:8113198
A;Accession: A53384
A;Molecule type: DNA
A;Residues: 1-377 <SWA>
A;Cross-references: GB:U04328; NID:g460144; PIDN:AAA17654.1; PID:g460145
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: D81241
A;Molecule type: DNA
A;Residues: 1-377 <TET>
A;Cross-references: GB:AE002366; GB:AE002098; NID:g7225284; PIDN:AAF40537.1; PID:g722
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0070

Query Match 46.2%; Score 48; DB 2; Length 377;
Best Local Similarity 52.9%; Pred. No. 8.8;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGSGVRGDFGSLAPRVA 17
| :| | ||| | | :|
Db 6 CITGTRADFGKLKPLLA 22

RESULT 6
S60758
siaA protein - Neisseria meningitidis
C;Species: Neisseria meningitidis
C;Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 08-Oct-1999
C;Accession: S60758
R;Edwards, U.; Mueller, A.; Hammerschmidt, S.; Gerardy-Schahn, R.; Frosch, M.
Mol. Microbiol. 14, 141-149, 1994
A;Title: Molecular analysis of the biosynthesis pathway of the alpha-2,8 polysialic a
A;Reference number: S60758; MUID:95131727; PMID:7830552
A;Accession: S60758
A;Status: preliminary; nucleic acid sequence not shown

```
A;Molecule type: DNA
A;Residues: 1-377 <EDW>
A;Cross-references: EMBL:M95053; NID:g520732; PIDN:AAA20475.1; PID:g520733
C;Genetics:
A;Gene: siaA

  Query Match             46.2%;  Score 48;  DB 2;  Length 377;
  Best Local Similarity   52.9%;  Pred. No. 8.8;
  Matches    9;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDFGSLAPRVA 17
              | :| | ||| | | :|
Db          6 CITGTRADFGKLKPLLA 22


RESULT 7
GNNY2F
genome polyprotein - foot-and-mouth disease virus A (strain A[10]61)
N;Contains: coat protein VP1; coat protein VP2; coat protein VP3; coat protein VP4; core
 protein p20a; nonstructural protein p20b; RNA-directed RNA polymerase (EC 2.7.7.48)
C;Species: Aphthovirus A (foot-and-mouth disease virus A)
C;Date: 17-Dec-1982 #sequence_revision 28-Aug-1985 #text_change 16-Jul-1999
C;Accession: A93508; A91491; S30753
R;Carroll, A.R.; Rowlands, D.J.; Clarke, B.E.
Nucleic Acids Res. 12, 2461-2472, 1984
A;Title: The complete nucleotide sequence of the RNA coding for the primary translation
A;Reference number: A93508; MUID:84169547; PMID:6324120
A;Accession: A93508
A;Molecule type: genomic RNA
A;Residues: 1-2333 <CAR>
A;Cross-references: GB:X00429
R;Boothroyd, J.C.; Harris, T.J.R.; Rowlands, D.J.; Lowe, P.A.
Gene 17, 153-161, 1982
A;Title: The nucleotide sequence of cDNA coding for the structural proteins of foot-and-
A;Reference number: A91491; MUID:82211814; PMID:6282711
A;Accession: A91491
A;Molecule type: genomic RNA
A;Residues: 115-395,'C',397-631,'L',633-1048 <BOO>
A;Cross-references: GB:V01130; NID:g61048; PIDN:CAA24361.1; PID:g1335402
R;Sangar, D.V.; Newton, S.E.; Rowlands, D.J.; Clarke, B.E.
Nucleic Acids Res. 15, 3305-3315, 1987
A;Title: All foot and mouth disease virus serotypes initiate protein synthesis at two se
A;Reference number: S30753; MUID:87203363; PMID:3033601
A;Accession: S30753
A;Molecule type: genomic RNA
A;Residues: 1-32 <SAN>
A;Cross-references: EMBL:M31575; NID:g210486; PIDN:AAA42655.1; PID:g210487
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: coat protein; core protein; genome-linked protein; nonstructural protein; nu
F;1-204/Product: nonstructural protein p20a #status predicted <NPA>
F;205-286/Product: coat protein VP4 #status predicted <VP4>
F;287-504/Product: coat protein VP2 #status predicted <VP2>
F;505-725/Product: coat protein VP3 #status predicted <VP3>
F;726-937/Product: coat protein VP1 #status predicted <VP1>
F;938-1578/Product: core protein p52 #status predicted <CPP>
F;1579-1601/Product: genome-linked protein VPg1 #status predicted <GL1>
F;1602-1625/Product: genome-linked protein VPg2 #status predicted <GL2>
F;1626-1649/Product: genome-linked protein VPg3 #status predicted <GL3>
F;1650-1863/Product: nonstructural protein p20b #status predicted <NPB>
F;1864-2333/Product: RNA-directed RNA polymerase #status predicted <RRP>

  Query Match             46.2%;  Score 48;  DB 1;  Length 2333;
  Best Local Similarity   71.4%;  Pred. No. 54;
  Matches   10;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 GDFGSLAPRVARQL 20
              || ||:| ||| ||
Db        869 GDLGSIAARVATQL 882


RESULT 8
E75631
iron ABC transporter, permease protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C;Accession: E75631
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: E75631
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-351 <WHI>
A;Cross-references: GB:AE001826; NID:g6460827; PIDN:AAF12554.1; PID:g6460850; TIGR:DR
A;Experimental source: strain R1
C;Genetics:
A;Gene: DRB0122
A;Map position: megaplasmid
A;Genome: plasmid
A;Note: plasmid MP1
C;Superfamily: vitamin B12 transport protein btuC

  Query Match             45.2%;  Score 47;  DB 2;  Length 351;
  Best Local Similarity   52.6%;  Pred. No. 12;
  Matches   10;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              |:|: |  | ||| :|| |
Db        275 GAGILGFVGLLAPHLARLL 293


RESULT 9
F96017
hypothetical protein SMb20699 [imported] - Sinorhizobium meliloti (strain 1021) magap
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: F96017
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing e
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: F96017
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-365 <KUR>
A;Cross-references: GB:AL591985; PIDN:CAC49806.1; PID:g15141294; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMb20699
A;Genome: plasmid
C;Superfamily: multidrug resistance protein A; lipoyl/biotin-binding homology

  Query Match             45.2%;  Score 47;  DB 2;  Length 365;
  Best Local Similarity   76.9%;  Pred. No. 12;
  Matches   10;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 VRGDFGSLAPRVA 17
              ||||  ||||:||
Db         63 VRGDVTSLAPKVA 75


RESULT 10
D70696
probable dtdp-glucose 4 - Mycobacterium tuberculosis (strain H37RV)
```

```
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C;Accession: D70696
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70696
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-326 <COL>
A;Cross-references: GB:Z80343; GB:AL123456; NID:g3261648; PIDN:CAB02463.1; PID:g1552867
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: epiB
C;Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
F;3-314/Domain: UDPglucose 4-epimerase homology <UDP>

  Query Match             44.7%;  Score 46.5;  DB 2;  Length 326;
  Best Local Similarity   57.9%;  Pred. No. 13;
  Matches   11;  Conservative    2;  Mismatches    3;  Indels    3;  Gaps    1;

Qy          4 GVR---GDFGSLAPRVARQ 19
              |||   | ||:| ||: ||
Db        179 GVRQKAGRFGALIPRLVRQ 197


RESULT 11
S61252
genome polyprotein - foot-and-mouth disease virus Asia (isolate Asia I Mandya, Karnataka
N;Alternate names: immunogenic polyprotein
C;Species: Aphthovirus Asia (foot-and-mouth disease virus Asia)
A;Variety: isolate Asia I Mandya, Karnataka
C;Date: 18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change 26-Aug-1999
C;Accession: S61252
R;Tulasiram, P.; Tyagi, M.; Suryanarayana, V.
submitted to the EMBL Data Library, June 1995
A;Description: Antigenic variation in foot and mouth disease virus type Asia I isolates
A;Reference number: S61252
A;Accession: S61252
A;Molecule type: mRNA
A;Residues: 1-109 <TUL>
A;Cross-references: EMBL:X88856; NID:g971407; PIDN:CAA61326.1; PID:g971408
A;Experimental source: isolate Asia I Mandya, Karnataka
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: coat protein; polyprotein; proteinase
F;1-91/Product: coat protein VP1 (fragment) #status predicted <MAT1>
F;92-107/Product: proteinase 2A #status predicted <MAT2>
F;108-109/Product: proteinase 2B (fragment) #status predicted <MAT3>

  Query Match             44.2%;  Score 46;  DB 2;  Length 109;
  Best Local Similarity   60.0%;  Pred. No. 5.3;
  Matches    9;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 RGDFGSLAPRVARQL 20
              |||  :|| |::|||
Db         23 RGDMAALAQRLSRQL 37


RESULT 12
S61253
genome polyprotein - foot-and-mouth disease virus Asia (isolate Asia I Nilgiri, Tamil Na
N;Alternate names: immunogenic polyprotein
N;Contains: coat protein VP1; proteinase 2A; proteinase 2B
C;Species: Aphthovirus Asia (foot-and-mouth disease virus Asia)
A;Variety: isolate Asia I Nilgiri, Tamil Nadu
C;Date: 18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change 26-Aug-1999
C;Accession: S61253
R;Tulasiram, P.; Tyagi, M.; Suryanarayana, V.
submitted to the EMBL Data Library, June 1995
A;Description: Antigenic variation in foot and mouth disease virus type Asia I isolat
A;Reference number: S61253
A;Accession: S61253
A;Molecule type: mRNA
A;Residues: 1-109 <TUL>
A;Cross-references: EMBL:X88857; NID:g971409; PIDN:CAA61327.1; PID:g971410
A;Experimental source: isolate Asia I Nilgiri, Tamil Nadu
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: coat protein; polyprotein; proteinase
F;1-91/Product: coat protein VP1 (fragment) #status predicted <MAT1>
F;92-107/Product: proteinase 2A #status predicted <MAT2>
F;108-109/Product: proteinase 2B (fragment) #status predicted <MAT3>

  Query Match             44.2%;  Score 46;  DB 2;  Length 109;
  Best Local Similarity   60.0%;  Pred. No. 5.3;
  Matches    9;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 RGDFGSLAPRVARQL 20
              |||  :|| |::|||
Db         23 RGDMAALAQRLSRQL 37


RESULT 13
S30395
ribosomal protein S2, cytosolic - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 13-Aug-1999
C;Accession: S30395; S50325; S30396
R;Barrio, R.; del Arco, A.; Cabrera, H.; Arribas, C.
Nucleic Acids Res. 21, 351, 1993
A;Title: Cloning and analysis of the S2 ribosomal protein cDNA from Drosophila.
A;Reference number: S30395; MUID:93181212; PMID:8441641
A;Accession: S30395
A;Molecule type: mRNA
A;Residues: 1-267 <BAR>
A;Cross-references: EMBL:X69120; NID:g288083; PIDN:CAA48872.1; PID:g288084
A;Note: the authors did not translate the codon for residue 267
R;Cramton, S.E.; Laski, F.A.
Genetics 137, 1039-1048, 1994
A;Title: string of pearls Encodes Drosophila ribosomal protein S2, has minute-like ch
A;Reference number: S50325; MUID:95073591; PMID:7982558
A;Accession: S50325
A;Molecule type: mRNA
A;Residues: 1-18,'GG',21-193,'K',195-267 <CRA>
A;Cross-references: EMBL:U01334; NID:g430711; PIDN:AAC34198.1; PID:g515971
C;Genetics:
A;Gene: FlyBase:sop
A;Cross-references: FlyBase:FBgn0004867
C;Superfamily: Escherichia coli ribosomal protein S5
C;Keywords: protein biosynthesis; ribosome

  Query Match             44.2%;  Score 46;  DB 2;  Length 267;
  Best Local Similarity   62.5%;  Pred. No. 13;
  Matches   10;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 SGVRGDFGSLAPRVAR 18
              || || |||  ||  |
Db          9 SGFRGGFGSRPPRGGR 24


RESULT 14
F90965
hypothetical protein ECs2694 [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C;Accession: F90965
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90965
A;Status: preliminary
```

A;Molecule type: DNA
A;Residues: 1-564 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB36117.1; PID:g13362162; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs2694

Query Match 44.2%; Score 46; DB 2; Length 564;
Best Local Similarity 80.0%; Pred. No. 27;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GVRGDFGSLA 13
||||||||::
Db 350 GVRGDFGSIS 359

RESULT 15
F85813
hypothetical protein Z3047 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: F85813
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: F85813
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-564 <STO>
A;Cross-references: GB:AE005174; NID:g12516062; PIDN:AAG56970.1; GSPDB:GN00145; UWGP:Z30
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z3047

Query Match 44.2%; Score 46; DB 2; Length 564;
Best Local Similarity 80.0%; Pred. No. 27;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GVRGDFGSLA 13
||||||||::
Db 350 GVRGDFGSIS 359

Search completed: October 9, 2003, 10:18:40
Job time : 21.6667 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:02:54 ; Search time 10.3333 Seconds
(without alignments)
91.019 Million cell updates/sec

Title: US-09-549-186B-7
Perfect score: 104
Sequence: 1 CGSGVRGDFGSLAPRVARQL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	95	91.3	2332	1	POLG_FMDVA	P03308	f genome po
2	67	64.4	216	1	POLG_FMDVC	P03309	foot-and-mo
3	50	48.1	230	1	POLG_FMDV5	P03307	foot-and-mo
4	50	48.1	2336	1	POLG_FMDVZ	P49303	f genome po
5	48	46.2	2333	1	POLG_FMDV1	P03306	f genome po
6	46	44.2	564	1	YEDQ_ECO57	Q8xb92	escherichia
7	46	44.2	564	1	YEDQ_ECOLI	P76330	escherichia
8	46	44.2	573	1	Y930_TREPA	O83900	treponema p
9	45.5	43.8	429	1	HRB1_YEAST	P38922	saccharomyc
10	44.5	42.8	1079	1	APCE_FREDI	P16566	fremyella d
11	44	42.3	379	1	CD62_METTH	O27636	methanobact
12	44	42.3	920	1	AD19_MOUSE	O35674	mus musculu
13	43	41.3	312	1	TRAD_HUMAN	Q15628	homo sapien
14	43	41.3	959	1	G2D1_HUMAN	Q9uhl9	h general t
15	42	40.4	309	1	FLIH_TREPA	O83416	treponema p
16	42	40.4	328	1	HXD1_HUMAN	Q9gzz0	homo sapien
17	42	40.4	486	1	K1C0_XENLA	P02537	xenopus lae
18	42	40.4	809	1	C48A_ARATH	P54609	arabidopsis
19	42	40.4	810	1	SYFB_SYNY3	P74296	synechocyst
20	42	40.4	980	1	CLV1_ARATH	Q9syq8	arabidopsis
21	41.5	39.9	217	1	SODM_DROME	Q00637	drosophila
22	41.5	39.9	933	1	PERT_CANFA	Q8hyb7	canis famil
23	41	39.4	775	1	SYFB_AQUAE	O67620	aquifex aeo
24	41	39.4	1355	1	CA21_RANCA	O42350	rana catesb
25	41	39.4	5147	1	FAT_DROME	P33450	drosophila
26	40.5	38.9	260	1	ETFB_CLOTS	P97089	clostridium
27	40.5	38.9	387	1	OMA2_NEIMC	P18194	neisseria m
28	40.5	38.9	392	1	OMA_NEIMB	Q51240	neisseria m
29	40.5	38.9	393	1	OMA1_NEIMC	P13415	neisseria m
30	40.5	38.9	395	1	OMA_NEIMA	P57041	neisseria m
31	40	38.5	170	1	RS9_STRCO	Q53875	streptomyce
32	40	38.5	231	1	CUT2_FUSSO	Q96ut0	fusarium so
33	40	38.5	231	1	CUT3_FUSSO	Q96us9	fusarium so
34	40	38.5	241	1	PYRF_THETH	P96076	thermus the
35	40	38.5	281	1	EF1D_HUMAN	P29692	homo sapien
36	40	38.5	281	1	WS22_MOUSE	Q9cy21	mus musculu
37	40	38.5	290	1	Y229_SYNY3	Q55702	synechocyst
38	40	38.5	330	1	PDXA_XYLFA	Q9pf39	xylella fas
39	40	38.5	348	1	ADH1_PICST	O00097	pichia stip
40	40	38.5	422	1	GFI1_HUMAN	Q99684	homo sapien
41	40	38.5	602	1	DNLI_AERPE	Q9yd18	aeropyrum p
42	40	38.5	654	1	YLH4_SCHPO	Q9hgq2	schizosacch
43	40	38.5	678	1	T2D5_MOUSE	Q62311	mus musculu
44	40	38.5	678	1	T2D5_RAT	Q63801	rattus norv
45	40	38.5	814	1	GUNE_CLOTM	P10477	clostridium

ALIGNMENTS

RESULT 1
POLG_FMDVA
ID POLG_FMDVA STANDARD; PRT; 2332 AA.
AC P03308; P03312; Q65038; Q65039; Q65040; Q65041; Q65042; Q65043;
AC Q65044; Q65045; Q65046; Q65047;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat
DE proteins VP1 TO VP4; Core proteins X, P14, P41, P19; Genome-linked
DE proteins VPG1 TO VPG3; Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48)].
OS Foot-and-mouth disease virus (strain A12) (Aphthovirus A) (FMDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Aphthovirus.
OX NCBI_TaxID=12114;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85211015; PubMed=2987518;
RA Robertson B.H., Grubman M.J., Weddell G.N., Moore D.M., Welsh J.D.,
RA Fischer T., Dowbenko D.J., Yansura D.G., Small B., Kleid D.G.;
RT "Nucleotide and amino acid sequence coding for polypeptides of
RT foot-and-mouth disease virus type A12.";
RL J. Virol. 54:651-660(1985).
RN [2]
RP SEQUENCE OF 1863-2332 FROM N.A.
RX MEDLINE=83225613; PubMed=6305004;
RA Robertson B.H., Morgan D.O., Moore D.M., Grubman M.J., Card J.,
RA Fischer T., Weddell G.N., Dowbenko D.J., Yansura D.G.;
RT "Identification of amino acid and nucleotide sequence of the
RT foot-and-mouth disease virus RNA polymerase.";
RL Virology 126:614-623(1983).
RN [3]
RP SEQUENCE OF 715-955 FROM N.A.
RX MEDLINE=82061853; PubMed=6272395;
RA Kleid D.G., Yansura D.G., Small B., Dowbenko D.J., Moore D.M.,
RA Grubman M.J., McKercher P.D., Morgan D.O., Robertson B.H.,
RA Bachrach H.L.;
RT "Cloned viral protein vaccine for foot-and-mouth disease: responses
RT in cattle and swine.";
RL Science 214:1125-1129(1981).
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC ----------

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M10975; AAA42593.1; -.
DR   EMBL; J02187; AAA42670.1; -.
DR   MEROPS; C03.008; -.
DR   InterPro; IPR004004; Calici_pol_hel.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   InterPro; IPR000605; RNA_helicase.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR001205; RNA_pol_P3D.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00073; rhv; 3.
DR   Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR   Pfam; PF00910; RNA_helicase; 1.
DR   PRINTS; PR00918; CALICIVIRUSNS.
DR   PRINTS; PR01542; FMDVP1COAT.
KW   Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;
KW   Transferase; Hydrolase; Thiol protease; Nonstructural protein;
KW   Myristate.
FT   CHAIN         1    200       NONSTRUCTURAL PROTEIN P20A.
FT   CHAIN       201    285       COAT PROTEIN VP4.
FT   CHAIN       286    503       COAT PROTEIN VP2.
FT   CHAIN       504    723       COAT PROTEIN VP3.
FT   CHAIN       724    937       COAT PROTEIN VP1.
FT   CHAIN       938    953       CORE PROTEIN X.
FT   CHAIN       954   1107       CORE PROTEIN P14.
FT   CHAIN      1108   1425       CORE PROTEIN P41.
FT   CHAIN      1426   1578       CORE PROTEIN P19.
FT   CHAIN      1579   1601       GENOME-LINKED PROTEIN VPG1.
FT   CHAIN      1602   1625       GENOME-LINKED PROTEIN VPG2.
FT   CHAIN      1626   1649       GENOME-LINKED PROTEIN VPG3.
FT   CHAIN      1650   1862       PROTEASE.
FT   CHAIN      1863   2332       RNA-DIRECTED RNA POLYMERASE.
FT   LIPID       201    201       MYRISTATE.
SQ   SEQUENCE   2332 AA;  259408 MW;  EE77DA739CBEDC6A CRC64;

  Query Match             91.3%;  Score 95;  DB 1;  Length 2332;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-06;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||||||||||||||||
Db        864 GSGVRGDFGSLAPRVARQL 882


RESULT 2
POLG_FMDVC
ID   POLG_FMDVC     STANDARD;      PRT;   216 AA.
AC   P03309;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Genome polyprotein [Contains: Coat protein VP1; Core protein P52]
DE   (Fragment).
OS   Foot-and-mouth disease virus (strain A24 Cruzeiro) (Aphthovirus A)
OS   (FMDV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12115;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83143292; PubMed=6298715;
RA   Makoff A.J., Paynter C.A., Rowlands D.J., Boothroyd J.C.;
RT   "Comparison of the amino acid sequence of the major immunogen from
RT   three serotypes of foot and mouth disease virus.";
RL   Nucleic Acids Res. 10:8285-8295(1982).
CC   -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC       EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC       VP3, AND VP4.
CC   -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; J02183; AAA42596.1; -.
DR   EMBL; A06733; CAA00589.1; -.
DR   PIR; A03911; A03911.
DR   HSSP; Q88571; 1TME.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
KW   Coat protein; Core protein; Polyprotein.
FT   NON_TER       1      1
FT   CHAIN         1    202       COAT PROTEIN VP1.
FT   CHAIN       203   >216       CORE PROTEIN P52.
FT   NON_TER     216    216
SQ   SEQUENCE   216 AA;  23889 MW;  501659FF031A1D85 CRC64;

  Query Match             64.4%;  Score 67;  DB 1;  Length 216;
  Best Local Similarity   73.7%;  Pred. No. 0.0036;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||| ||| |||| || :||
Db        129 GSGRRGDMGSLAARVVKQL 147


RESULT 3
POLG_FMDV5
ID   POLG_FMDV5     STANDARD;      PRT;   230 AA.
AC   P03307;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Genome polyprotein [Contains: Coat proteins VP3, VP1; Core protein
DE   P52] (Fragment).
OS   Foot-and-mouth disease virus (strain A5) (Aphthovirus A) (FMDV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12113;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84028562; PubMed=6194987;
RA   Beck E., Feil G., Strohmaier K.;
RT   "The molecular basis of the antigenic variation of foot-and-mouth
RT   disease virus.";
RL   EMBO J. 2:555-559(1983).
CC   -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC       EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC       VP3, AND VP4.
CC   -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; V01135; CAA24365.1; ALT_INIT.
DR   EMBL; V01135; CAA24366.1; ALT_SEQ.
DR   PIR; A03909; A03909.
DR   HSSP; Q88571; 1TME.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
```

```
DR   PRINTS; PR01542; FMDVP1COAT.
KW   Coat protein; Core protein; Polyprotein.
FT   NON_TER       1      1
FT   CHAIN        <1      4       COAT PROTEIN VP3.
FT   CHAIN         5    216       COAT PROTEIN VP1.
FT   CHAIN       217   >230       CORE PROTEIN P52.
FT   NON_TER     230    230
SQ   SEQUENCE   230 AA;  25369 MW;  EA93A190F4CC1608 CRC64;

  Query Match             48.1%;  Score 50;  DB 1;  Length 230;
  Best Local Similarity   57.9%;  Pred. No. 1.6;
  Matches   11;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              |   ||| || | | |:||
Db        143 GGPRRGDMGSAAARAAKQL 161


RESULT 4
POLG_FMDVZ
ID   POLG_FMDVZ     STANDARD;      PRT;  2336 AA.
AC   P49303;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Genome polyprotein [Contains: Nonstructural protein P20A; Coat-
DE   proteins VP1 TO VP4; Core proteins X, P14, P41, P19; Genome-linked
DE   proteins VPG1 TO VPG3; Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE   (P3C); RNA-directed RNA polymerase (EC 2.7.7.48)].
OS   Foot-and-mouth disease virus (strain A22/550 Azerbaijan 65)
OS   (Aphthovirus A) (FMDV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=73481;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sosnovtsev S.V., Onischenko A.M., Petrov N.A., Kalashnikova T.I.,
RA   Mamaeva N.V., Drygin V.Y., Perevozchikova N.A., Vasilenko S.K.;
RL   Submitted (AUG-1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the
CC       poliovirus polyprotein. In other picornavirus reactions Glu may be
CC       substituted for Gln, and Ser or Thr for Gly.
CC   -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC       {RNA}(N).
CC   -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC       EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC       VP3, AND VP4.
CC   -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC   -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X74812; CAA52812.1; -.
DR   PIR; S37077; S37077.
DR   HSSP; Q88571; 1TME.
DR   MEROPS; C03.008; -.
DR   MEROPS; C28.001; -.
DR   InterPro; IPR004004; Calici_pol_hel.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   InterPro; IPR000605; RNA_helicase.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR001205; RNA_pol_P3D.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00073; rhv; 3.
DR   Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR   Pfam; PF00910; RNA_helicase; 1.
DR   PRINTS; PR00918; CALICVIRUSNS.
DR   PRINTS; PR01542; FMDVP1COAT.
KW   Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;
KW   Transferase; Hydrolase; Thiol protease; Nonstructural protein;
KW   Myristate.
FT   CHAIN         1    201       NONSTRUCTURAL PROTEIN P20A.
FT   CHAIN       202    286       COAT PROTEIN VP4.
FT   CHAIN       287    504       COAT PROTEIN VP2.
FT   CHAIN       505    724       COAT PROTEIN VP3.
FT   CHAIN       725    938       COAT PROTEIN VP1.
FT   CHAIN       939    954       CORE PROTEIN X.
FT   CHAIN       955   1108       CORE PROTEIN P14.
FT   CHAIN      1109   1426       CORE PROTEIN P41.
FT   CHAIN      1427   1579       CORE PROTEIN P19.
FT   CHAIN      1580   1602       GENOME-LINKED PROTEIN VPG1.
FT   CHAIN      1603   1626       GENOME-LINKED PROTEIN VPG2.
FT   CHAIN      1627   1650       GENOME-LINKED PROTEIN VPG3.
FT   CHAIN      1651   1863       PROTEASE.
FT   CHAIN      1864   2336       RNA-DIRECTED RNA POLYMERASE.
FT   LIPID       202    202       MYRISTATE (BY SIMILARITY).
SQ   SEQUENCE   2336 AA;  259982 MW;  15AC2AB022B5B954 CRC64;

  Query Match             48.1%;  Score 50;  DB 1;  Length 2336;
  Best Local Similarity   63.2%;  Pred. No. 15;
  Matches   12;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              | | |||   || ||| ||
Db        865 GMGRRGDLEPLAARVAAQL 883


RESULT 5
POLG_FMDV1
ID   POLG_FMDV1     STANDARD;      PRT;  2333 AA.
AC   P03306; Q64768; Q84750; Q84751; Q84752; Q84753; Q84754; Q84760;
AC   Q84761; Q84762; Q84763; Q84764; Q84765; Q84766; Q84767; Q84768;
AC   Q84769; Q89824;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Genome polyprotein [Contains: Nonstructural protein P20A; Coat
DE   proteins VP1 TO VP4; Core protein P52; Genome-linked proteins VPG1 TO
DE   VPG3; Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed
DE   RNA polymerase P56A (EC 2.7.7.48)].
OS   Foot-and-mouth disease virus (strain A10-61) (Aphthovirus A) (FMDV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12112;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84169547; PubMed=6324120;
RA   Carroll A.R., Rowlands D.J., Clarke B.E.;
RT   "The complete nucleotide sequence of the RNA coding for the primary
RT   translation product of foot and mouth disease virus.";
RL   Nucleic Acids Res. 12:2461-2472(1984).
RN   [2]
RP   SEQUENCE OF 115-1048 FROM N.A.
RX   MEDLINE=82211814; PubMed=6282711;
RA   Boothroyd J.C., Harris T.J.R., Rowlands D.J., Lowe P.A.;
RT   "The nucleotide sequence of cDNA coding for the structural proteins
RT   of foot-and-mouth disease virus.";
RL   Gene 17:153-161(1982).
CC   -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the
CC       poliovirus polyprotein. In other picornavirus reactions Glu may be
CC       substituted for Gln, and Ser or Thr for Gly.
CC   -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC       {RNA}(N).
CC   -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC       EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC       VP3, AND VP4.
CC   -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC   -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC   ---------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; V01130; CAA24361.1; -.
DR   EMBL; X00429; CAA25127.1; -.
DR   MEROPS; C03.008; -.
DR   MEROPS; C28.001; -.
DR   InterPro; IPR004004; Calici_pol_hel.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   InterPro; IPR000605; RNA_helicase.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR001205; RNA_pol_P3D.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00073; rhv; 3.
DR   Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR   Pfam; PF00910; RNA_helicase; 1.
DR   PRINTS; PR00918; CALICIVIRUSNS.
DR   PRINTS; PR01542; FMDVP1COAT.
KW   Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;
KW   Transferase; Hydrolase; Thiol protease; Nonstructural protein;
KW   Myristate.
FT   CHAIN         1    201       NONSTRUCTURAL PROTEIN P20A.
FT   CHAIN       202    286       COAT PROTEIN VP4.
FT   CHAIN       287    504       COAT PROTEIN VP2.
FT   CHAIN       505    725       COAT PROTEIN VP3.
FT   CHAIN       726    937       COAT PROTEIN VP1.
FT   CHAIN       938   1578       CORE PROTEIN P52.
FT   CHAIN      1579   1601       GENOME-LINKED PROTEIN VPG1.
FT   CHAIN      1602   1625       GENOME-LINKED PROTEIN VPG2.
FT   CHAIN      1626   1649       GENOME-LINKED PROTEIN VPG3.
FT   CHAIN      1650   1863       PROTEASE P20B.
FT   CHAIN      1864   2333       RNA-DIRECTED RNA POLYMERASE P56A.
FT   LIPID       202    202       MYRISTATE.
FT   CONFLICT    396    396       S -> C (IN REF. 2).
FT   CONFLICT    632    632       P -> L (IN REF. 2).
SQ   SEQUENCE   2333 AA;  259645 MW;  4FC667DCC521BC60 CRC64;

  Query Match             46.2%;  Score 48;  DB 1;  Length 2333;
  Best Local Similarity   71.4%;  Pred. No. 30;
  Matches   10;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 GDFGSLAPRVARQL 20
              || ||:| ||| ||
Db        869 GDLGSIAARVATQL 882


RESULT 6
YEDQ_ECO57
ID   YEDQ_ECO57     STANDARD;      PRT;   564 AA.
AC   Q8XB92;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein yedQ.
GN   YEDQ OR Z3047 OR ECS2694.
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: BELONGS TO THE YAIC / YFIN (E.COLI), YHCK (B.SUBTILIS)
CC       FAMILY.
CC   -!- SIMILARITY: Contains 1 GGDEF domain.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AE005417; AAG56970.1; -.
DR   EMBL; AP002559; BAB36117.1; -.
DR   PIR; F90965; F90965.
DR   InterPro; IPR000160; GGDEF.
DR   Pfam; PF00990; GGDEF; 1.
DR   SMART; SM00267; DUF1; 1.
DR   TIGRFAMs; TIGR00254; GGDEF; 1.
DR   PROSITE; PS50887; GGDEF; 1.
KW   Hypothetical protein; Transmembrane; Complete proteome.
FT   TRANSMEM     20     40       POTENTIAL.
FT   TRANSMEM    360    380       POTENTIAL.
FT   DOMAIN      428    563       GGDEF.
SQ   SEQUENCE   564 AA;  64287 MW;  BEBC2286ADBAECB0 CRC64;

  Query Match             44.2%;  Score 46;  DB 1;  Length 564;
  Best Local Similarity   80.0%;  Pred. No. 15;
  Matches    8;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GVRGDFGSLA 13
              ||||||||::
Db        350 GVRGDFGSIS 359


RESULT 7
YEDQ_ECOLI
ID   YEDQ_ECOLI     STANDARD;      PRT;   564 AA.
AC   P76330; P94746;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein yedQ.
GN   YEDQ OR B1956.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
```

```
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251358; PubMed=9097040;
RA   Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA   Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
RA   Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
RA   Nashimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA   Sivasundaram S., Tagami H., Takeda J., Takemoto K., Wada C.,
RA   Yamamoto Y., Horiuchi T.;
RT   "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 40.1-50.0 min region on the linkage map.";
RL   DNA Res. 3:379-392(1996).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: BELONGS TO THE YAIC / YFIN (E.COLI), YHCK (B.SUBTILIS)
CC       FAMILY.
CC   -!- SIMILARITY: Contains 1 GGDEF domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000287; AAC75022.1; ALT_INIT.
DR   EMBL; D90835; BAA15784.1; -.
DR   EcoGene; EG14040; yedQ.
DR   InterPro; IPR000160; GGDEF.
DR   Pfam; PF00990; GGDEF; 1.
DR   SMART; SM00267; DUF1; 1.
DR   TIGRFAMs; TIGR00254; GGDEF; 1.
DR   PROSITE; PS50887; GGDEF; 1.
KW   Hypothetical protein; Transmembrane; Complete proteome.
FT   TRANSMEM     20     40       POTENTIAL.
FT   TRANSMEM    360    380       POTENTIAL.
FT   DOMAIN      428    563       GGDEF.
SQ   SEQUENCE   564 AA;  64283 MW;  05FB02C1BE2A8938 CRC64;

  Query Match             44.2%;  Score 46;  DB 1;  Length 564;
  Best Local Similarity   80.0%;  Pred. No. 15;
  Matches    8;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GVRGDFGSLA 13
              ||||||||::
Db        350 GVRGDFGSIS 359


RESULT 8
Y930_TREPA
ID   Y930_TREPA     STANDARD;      PRT;   573 AA.
AC   O83900;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein TP0930.
GN   TP0930.
OS   Treponema pallidum.
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX   NCBI_TaxID=160;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Nichols;
RX   MEDLINE=98332770; PubMed=9665876;
RA   Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA   Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA   Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA   Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA   McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA   Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA   Venter J.C.;
RT   "Complete genome sequence of Treponema pallidum, the syphilis
RT   spirochete.";
RL   Science 281:375-388(1998).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001261; AAC65889.1; -.
DR   PIR; C71264; C71264.
DR   TIGR; TP0930; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   573 AA;  64723 MW;  4A085F7C7B612E6A CRC64;

  Query Match             44.2%;  Score 46;  DB 1;  Length 573;
  Best Local Similarity   56.2%;  Pred. No. 16;
  Matches    9;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 SGVRGDFGSLAPRVAR 18
              :|||||  || ||: :
Db         12 AGVRGDVPSLPPRLIK 27


RESULT 9
HRB1_YEAST
ID   HRB1_YEAST     STANDARD;      PRT;   429 AA.
AC   P38922;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   HRB1 protein (TOM34 protein).
GN   HRB1 OR TOM34 OR YNL004W OR N2009.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GRF88;
RX   MEDLINE=94005822; PubMed=8402262;
RA   Lalo D., Stettler S., Mariotte S., Slonimski P.P., Thuriaux P.;
RT   "Two yeast chromosomes are related by a fossil duplication of their
RT   centromeric regions.";
RL   C. R. Acad. Sci., III, Sci. Vie 316:367-373(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GRF88;
RX   MEDLINE=95028151; PubMed=7941739;
RA   Lalo D., Stettler S., Mariotte S., Gendreau E., Thuriaux P.;
RT   "Organization of the centromeric region of chromosome XIV in
RT   Saccharomyces cerevisiae.";
RL   Yeast 10:523-533(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / FY1679;
RX   MEDLINE=95076713; PubMed=7985421;
RA   Verhasselt P., Aert R., Voet M., Volckaert G.;
RT   "Nucleotide sequence analysis of an 8887 bp region of the left arm of
RT   yeast chromosome XIV, encompassing the centromere sequence.";
RL   Yeast 10:945-951(1994).
CC   -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC   -!- SIMILARITY: STRONG, TO YEAST GBP2.
CC   -!- SIMILARITY: Contains 3 RNA recognition motif (RRM) domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U02536; AAA64803.1; -.
DR   EMBL; X77114; CAA54378.1; -.
DR   EMBL; Z71280; CAA95863.1; -.
DR   PIR; S45459; S45459.
DR   HSSP; P09651; 1HA1.
DR   SGD; S0004949; HRB1.
DR   GO; GO:0005737; C:cytoplasm; IDA.
DR   GO; GO:0030530; C:heterogeneous nuclear ribonucleoprotein com. . .; NAS.
DR   GO; GO:0006606; P:protein-nucleus import; IPI.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   Pfam; PF00076; rrm; 3.
DR   SMART; SM00360; RRM; 3.
DR   PROSITE; PS50102; RRM; 3.
DR   PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW   RNA-binding; Nuclear protein; Repeat.
FT   DOMAIN      136    212       RNA-BINDING (RRM) 1.
FT   DOMAIN      236    313       RNA-BINDING (RRM) 2.
FT   DOMAIN      351    428       RNA-BINDING (RRM) 3.
SQ   SEQUENCE   429 AA;  49141 MW;  DC0F732EFA43EE89 CRC64;

  Query Match             43.8%;  Score 45.5;  DB 1;  Length 429;
  Best Local Similarity   55.6%;  Pred. No. 14;
  Matches   10;  Conservative    3;  Mismatches    2;  Indels    3;  Gaps    1;

Qy          3 SGVRGDFGSLAPRVARQL 20
              || |||:|   | :||:|
Db        108 SGARGDYG---PLLAREL 122


RESULT 10
APCE_FREDI
ID   APCE_FREDI     STANDARD;      PRT;  1079 AA.
AC   P16566;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   Phycobilisome 120 kDa linker polypeptide, core (L-CM 92) (Core-
DE   membrane linker protein).
GN   APCE.
OS   Fremyella diplosiphon (Calothrix PCC 7601).
OC   Bacteria; Cyanobacteria; Nostocales; Microchaetaceae; Fremyella.
OX   NCBI_TaxID=1197;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90192765; PubMed=2107546;
RA   Houmard J., Capuano V., Colombano M.V., Coursin T., de Marsac N.;
RT   "Molecular characterization of the terminal energy acceptor of
RT   cyanobacterial phycobilisomes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:2152-2156(1990).
CC   -!- FUNCTION: THIS PROTEIN IS POSTULATED TO ACT BOTH AS TERMINAL
CC       ENERGY ACCEPTOR (BY ITS PHYCOBILIN-LIKE DOMAINS) AND AS A LINKER
CC       POLYPEPTIDE (BY ITS REPEATS AND ARMS) THAT STABILIZES THE
CC       PHYCOBILISOME CORE ARCHITECTURE.
CC   -!- SUBUNIT: PHYCOBILISOMES OF THIS ORGANISM ARE COMPOSED OF A TWO
CC       CYLINDER CORE, FROM WHICH SIX RODS RADIATE. THE CORE IS MAINLY
CC       COMPOSED OF ALLOPHYCOCYANIN ALPHA AND BETA CHAINS, AND OF THREE
CC       MINOR COMPONENTS: THE ALLOPHYCOCYANIN ALPHA-B CHAIN, A 18.3 kDa
CC       POLYPEPTIDE, AND THE ANCHOR POLYPEPTIDE LCM.
CC   -!- SUBCELLULAR LOCATION: ANCHORS THE PHYCOBILISOME PERPENDICULARLY
CC       TO THE STROMAL SURFACE OF THE THYLAKOID MEMBRANE.
CC   -!- SIMILARITY: THE REPEATED DOMAINS ARE SIMILAR TO THE N-TERMINAL
CC       REGIONS OF PHYCOCYANIN ROD LINKER POLYPEPTIDES.
CC   -!- SIMILARITY: THE PHYCOBILIN-LIKE DOMAINS ARE SIMILAR TO PHYCOBILINS
CC       FROM VARIOUS SPECIES.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M20806; AAA24873.1; -.
DR   PIR; A35088; A35088.
DR   HSSP; P00318; 1B33.
DR   InterPro; IPR001297; PBS_linker_poly.
DR   InterPro; IPR001659; Phycobilisome.
DR   Pfam; PF00427; PBS_linker_poly; 4.
DR   Pfam; PF00502; Phycobilisome; 1.
DR   ProDom; PD000852; PBS_linker_poly; 4.
DR   ProDom; PD000340; Phycobilisome; 2.
KW   Phycobilisome; Electron transport; Photosynthesis; Repeat.
FT   INIT_MET      0      0
FT   DOMAIN       17     75       PHYCOBILIN-LIKE 1.
FT   DOMAIN       76    143       PHYCOBILIN-LIKE LOOP.
FT   DOMAIN      144    236       PHYCOBILIN-LIKE 2.
FT   DOMAIN      237    284       ARM 1 (SPACING SEQUENCE).
FT   REPEAT      285    409       I.
FT   DOMAIN      410    546       ARM 2 (SPACING SEQUENCE).
FT   REPEAT      547    669       II.
FT   DOMAIN      670    743       ARM 3 (SPACING SEQUENCE).
FT   REPEAT      744    869       III.
FT   DOMAIN      870    953       ARM 4 (SPACING SEQUENCE).
FT   REPEAT      954   1079       IV.
SQ   SEQUENCE  1079 AA;  120325 MW;  87FE38F232BFCF82 CRC64;

  Query Match             42.8%;  Score 44.5;  DB 1;  Length 1079;
  Best Local Similarity   62.5%;  Pred. No. 49;
  Matches   10;  Conservative    2;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          4 GVRGDF-GSLAPRVAR 18
              | ||:| ||| |:| |
Db        504 GARGEFPGSLGPKVFR 519


RESULT 11
CD62_METTH
ID   CD62_METTH     STANDARD;      PRT;   379 AA.
AC   O27636;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cell division control protein 6 homolog 2 (CDC6 homolog 2).
GN   CDC6-2 OR MTH1599.
OS   Methanobacterium thermoautotrophicum.
OC   Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC   Methanobacteriaceae; Methanothermobacter.
OX   NCBI_TaxID=187420;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Delta H;
RX   MEDLINE=98037514; PubMed=9371463;
RA   Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., Dubois J.,
RA   Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA   Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA   Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA   Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA   McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA   Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT   "Complete genome sequence of Methanobacterium thermoautotrophicum
RT   deltaH: functional analysis and comparative genomics.";
RL   J. Bacteriol. 179:7135-7155(1997).
RN   [2]
RP   PHOSPHORYLATION, AND MUTAGENESIS OF LYS-71 AND ASP-149.
RX   MEDLINE=21405752; PubMed=11514535;
RA   Grabowski B., Kelman Z.;
RT   "Autophosphorylation of archaeal Cdc6 homologues is regulated by
RT   DNA.";
RL   J. Bacteriol. 183:5459-5464(2001).
CC   -!- FUNCTION: Might be involved in the initiation of DNA replication.
CC   -!- INDUCTION: Both single-stranded DNA and double-stranded DNA
CC       inhibit the phosyphorylation reaction.
```

```
CC   -!- PTM: Autophosphorylated on a serine.
CC   -!- SIMILARITY: BELONGS TO THE CDC6/CDC18 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000919; AAB86072.1; -.
DR   PIR; D69080; D69080.
DR   PhosSite; O27636; -.
DR   HAMAP; MF_01407; -; 1.
KW   DNA replication; Phosphorylation; ATP-binding; Complete proteome.
FT   NP_BIND      65     72       ATP.
FT   MUTAGEN      71     71       K->E: NO LABELING WITH ATP.
FT   MUTAGEN     148    148       D->N: RETAINS SOME ABILITY TO BE LABELED.
SQ   SEQUENCE   379 AA;  43301 MW;  48692C544BF94D01 CRC64;

  Query Match             42.3%;  Score 44;  DB 1;  Length 379;
  Best Local Similarity   50.0%;  Pred. No. 21;
  Matches    9;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQ 19
              | ||||:   | || :|:
Db        354 GKGVRGNSRILIPRFSRE 371


RESULT 12
AD19_MOUSE
ID   AD19_MOUSE     STANDARD;      PRT;   920 AA.
AC   O35674;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   ADAM 19 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE   domain 19) (Meltrin beta).
GN   ADAM19 OR MLTNB.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Myoblasts;
RX   MEDLINE=98129833; PubMed=9461614;
RA   Inoue D., Reid M.S., Lum L., Kraetzschmar J., Weskamp G., Myung Y.M.,
RA   Baron R., Blobel C.P.;
RT   "Cloning and initial characterization of mouse meltrin beta and
RT   analysis of the expression of four metalloprotease-disintegrins in
RT   bone cells.";
RL   J. Biol. Chem. 273:4180-4187(1998).
RN   [2]
RP   SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGE.
RC   TISSUE=Myoblasts;
RX   MEDLINE=98288128; PubMed=9622634;
RA   Kurisaki T., Masuda A., Osumi N., Nabeshima Y.-I., Fujisawa-Sehara A.;
RT   "Spatially- and temporally-restricted expression of meltrin alpha
RT   (ADAM12) and beta (ADAM19) in mouse embryo.";
RL   Mech. Dev. 73:211-215(1998).
RN   [3]
RP   SEQUENCE OF 429-578 FROM N.A.
RC   TISSUE=Embryonic fibroblast;
RX   MEDLINE=96026308; PubMed=7566181;
RA   Yagami-Hiromasa T., Sato T., Kurisaki T., Kamijo K., Nabeshima Y.-I.,
RA   Fujisawa-Sehara A.;
RT   "A metalloprotease-disintegrin participating in myoblast fusion.";
RL   Nature 377:652-656(1995).
RN   [4]
RP   FUNCTION.
RX   MEDLINE=21153759; PubMed=11116142;
RA   Shirakabe K., Wakatsuki S., Kurisaki T., Fujisawa-Sehara A.;
RT   "Roles of Meltrin beta /ADAM19 in the processing of neuregulin.";
RL   J. Biol. Chem. 276:9352-9358(2001).
CC   -!- FUNCTION: PARTICIPATES IN THE PROTEOLYTIC PROCESSING OF BETA-TYPE
CC       NEUREGULIN ISOFORMS WHICH ARE INVOLVED IN NEUROGENESIS AND
CC       SYNAPTOGENESIS, SUGGESTING A REGULATORY ROLE IN GLIAL CELL. ALSO
CC       CLEAVES ALPHA-2 MACROGLOBULIN. MAY BE INVOLVED IN OSTEOBLAST
CC       DIFFERENCIATION AND/OR OSTEOBLAST ACTIVITY IN BONE (BY
CC       SIMILARITY).
CC   -!- COFACTOR: BINDS ONE ZINC ION (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED, WITH THE HIGHEST EXPRESSION
CC       IN BONE, HEART AND LUNG, FOLLOWED BY BRAIN AND SPLEEN AND
CC       RELATIVELY LOW EXPRESSION IN LIVER, SKELETAL MUSCLE, KIDNEY AND
CC       TESTIS. IN BONE, PRIMARILY EXPRESSED IN CELL OF THE OSTEOBLAST
CC       LINEAGE AND NOT DETECTED IN MATURE OSTEOCLASTS.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED IN THE HEART AND IN THE TAIL BUD AT
CC       8.0 DPC, AND THEN IN THE CRANIAL AND DORSAL ROOT GANGLIA. ALSO
CC       EXPRESSED WEAKLY AND TRANSIENTLY IN THE INTESTINE, LUNG AND IN
CC       BONE MARROW.
CC   -!- INDUCTION: BY CALCITRIOL AND DURING OSTEOBLAST DIFFERENTIATION.
CC   -!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: Belongs to peptidase family M12B.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 1 disintegrin domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF019887; AAC40037.1; -.
DR   EMBL; D50410; BAA18923.2; -.
DR   HSSP; P18619; 1FVL.
DR   MEROPS; M12.214; -.
DR   MGD; MGI:105377; Adam19.
DR   InterPro; IPR006586; ADAM_cysteine.
DR   InterPro; IPR001762; Disintegrin.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR001818; Matrixin.
DR   InterPro; IPR002870; Pep_M12B_propep.
DR   InterPro; IPR001590; Reprolysin.
DR   InterPro; IPR006025; Zn_MTpeptdse.
DR   Pfam; PF00200; disintegrin; 1.
DR   Pfam; PF01562; Pep_M12B_propep; 1.
DR   Pfam; PF01421; Reprolysin; 1.
DR   PRINTS; PR00289; DISINTEGRIN.
DR   ProDom; PD000664; Disintegrin; 1.
DR   SMART; SM00608; ACR; 1.
DR   SMART; SM00050; DISIN; 1.
DR   PROSITE; PS50215; ADAM_MEPRO; 1.
DR   PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
DR   PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR   PROSITE; PS50214; DISINTEGRIN_2; 1.
DR   PROSITE; PS00022; EGF_1; FALSE_NEG.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW   Transmembrane; EGF-like domain; SH3-binding.
FT   SIGNAL       1     26       POTENTIAL.
FT   PROPEP      27    204       BY SIMILARITY.
FT   CHAIN      205    920       ADAM 19.
FT   DOMAIN      27    703       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   704    724       POTENTIAL.
FT   DOMAIN     725    920       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN     211    409       METALLOPROTEASE.
FT   DOMAIN     417    503       DISINTEGRIN-LIKE.
FT   DOMAIN     435    438       POLY-GLU.
FT   DOMAIN     504    653       CYS-RICH.
```

```
FT   DOMAIN      654    686       EGF-LIKE.
FT   DOMAIN      616    621       POLY-GLU.
FT   SITE        835    841       SH3-BINDING (POTENTIAL).
FT   SITE        840    846       SH3-BINDING (POTENTIAL).
FT   SITE        133    133       CYSTEINE SWITCH (POTENTIAL).
FT   METAL       346    346       ZINC (CATALYTIC).
FT   ACT_SITE    347    347
FT   METAL       350    350       ZINC (CATALYTIC).
FT   METAL       356    356       ZINC (CATALYTIC).
FT   DISULFID    321    404       BY SIMILARITY.
FT   DISULFID    361    388       BY SIMILARITY.
FT   DISULFID    466    488       POTENTIAL.
FT   DISULFID    658    668       BY SIMILARITY.
FT   DISULFID    662    674       BY SIMILARITY.
FT   DISULFID    676    685       BY SIMILARITY.
FT   CARBOHYD    145    145       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    445    445       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    448    448       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    649    649       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   920 AA;  100860 MW;  7094FDD4EE547382 CRC64;

  Query Match             42.3%;  Score 44;  DB 1;  Length 920;
  Best Local Similarity   38.9%;  Pred. No. 50;
  Matches    7;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDFGSLAPRVAR 18
              || |: | :   :|| |:
Db        555 CGKGLNGQYRKCSPRDAK 572


RESULT 13
TRAD_HUMAN
ID   TRAD_HUMAN     STANDARD;      PRT;   312 AA.
AC   Q15628;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Tumor necrosis factor receptor type 1 associated DEATH domain protein
DE   (TNFR1-associated DEATH domain protein) (TNFRSF1A-associated via death
DE   domain).
GN   TRADD.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95277836; PubMed=7758105;
RA   Hsu H., Xiong J., Goeddel D.V.;
RT   "The TNF receptor 1-associated protein TRADD signals cell death and
RT   NF-kappa B activation.";
RL   Cell 81:495-504(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Scheuerpflug C.G., Dechant M., Fellenberg J., Ewerbeck V.,
RA   Debatin K.M.;
RT   "Sequence, genomic organisation, and mutation analysis of the human
RT   TRADD gene in childhood B- and T-lineage acute lymphoblastic leukemia
RT   and ALPS.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreas;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   INTERACTION WITH RIPK1.
RX   MEDLINE=96200892; PubMed=8612133;
RA   Hsu H., Huang J., Shu H.-B., Baichwal V.R., Goeddel D.V.;
RT   "TNF-dependent recruitment of the protein kinase RIP to the TNF
RT   receptor-1 signaling complex.";
RL   Immunity 4:387-396(1996).
RN   [5]
RP   STRUCTURE BY NMR OF 1-179.
RX   PubMed=10911999;
RA   Tsao D.H., McDonagh T., Telliez J.B., Hsu S., Malakian K., Xu G.Y.,
RA   Lin L.L.;
RT   "Solution structure of N-TRADD and characterization of the interaction
RT   of N-TRADD and C-TRAF2, a key step in the TNFR1 signaling pathway.";
RL   Mol. Cell 5:1051-1057(2000).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 1-179 IN COMPLEX WITH
RP   TRAF2.
RX   PubMed=10892748;
RA   Park Y.C., Ye H., Hsia C., Segal D., Rich R.L., Liou H.C.,
RA   Myszka D.G., Wu H.;
RT   "A novel mechanism of TRAF signaling revealed by structural and
RT   functional analyses of the TRADD-TRAF2 interaction.";
RL   Cell 101:777-787(2000).
CC   -!- FUNCTION: Adapter molecule for TNFRSF1A/TNFR1 that specifically
CC       associates with the cytoplasmic domain of activated TNFRSF1A/TNFR1
CC       mediating its interaction with FADD. Overexpression of TRADD leads
CC       to two major TNF-induced responses, apoptosis and activation of
CC       NF-kappa-B.
CC   -!- SUBUNIT: Heterodimer with TNFRSF1A/TNFR1. Binds to TRAF2.
CC       Interacts with TRAF1, FADD and RIPK1.
CC   -!- TISSUE SPECIFICITY: Found in all examined tissues.
CC   -!- DOMAIN: Requires the intact DEATH domain to associate with
CC       TNFRSF1A/TNFR1.
CC   -!- SIMILARITY: Contains 1 death domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L41690; AAA98482.1; ALT_INIT.
DR   EMBL; AJ311614; CAC38018.2; -.
DR   EMBL; AJ311615; CAC38018.2; JOINED.
DR   EMBL; AJ311616; CAC38018.2; JOINED.
DR   EMBL; BC004491; AAH04491.1; -.
DR   PIR; A56911; A56911.
DR   PDB; 1F2H; 24-MAY-00.
DR   PDB; 1F3V; 06-JUN-00.
DR   GeneW; HGNC:12030; TRADD.
DR   MIM; 603500; -.
DR   GO; GO:0006917; P:induction of apoptosis; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR000488; Death.
DR   Pfam; PF00531; death; 1.
DR   SMART; SM00005; DEATH; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Apoptosis; 3D-structure.
FT   DOMAIN      179    289       DEATH.
```

```
FT   DOMAIN      192    198       POLY-PRO.
SQ   SEQUENCE   312 AA;  34247 MW;  5645D7E63E5FF05A CRC64;

  Query Match             41.3%;  Score 43;  DB 1;  Length 312;
  Best Local Similarity   61.5%;  Pred. No. 25;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDFGSLA 13
              |||| ||  | :|
Db        164 CGSGARGGDGEVA 176


RESULT 14
G2D1_HUMAN
ID   G2D1_HUMAN     STANDARD;      PRT;   959 AA.
AC   Q9UHL9; O95444; Q8WVC4; Q9UHK8; Q9UI91;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   General transcription factor II-I repeat domain-containing protein 1
DE   (GTF2I repeat domain containing protein 1) (Muscle TFII-I repeat
DE   domain-containing protein 1) (General transcription factor III) (Slow-
DE   muscle-fiber enhancer binding protein) (USE B1 binding protein)
DE   (MusTRD1/BEN) (Williams-Beuren syndrome chromosome region 11 protein).
GN   GTF2IRD1 OR MUSTRD1 OR GTF3 OR CREAM1 OR RBAP2 OR WBSCR11 OR WBSCR12.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Muscle;
RX   MEDLINE=98449952; PubMed=9774679;
RA   O'Mahoney J.V., Guven K.L., Lin J., Joya J.E., Robinson C.S.,
RA   Wade R.P., Hardeman E.C.;
RT   "Identification of a novel slow-muscle-fiber enhancer binding protein,
RT   MUSTRD1.";
RL   Mol. Cell. Biol. 18:6641-6652(1998).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Colon carcinoma;
RX   MEDLINE=99216421; PubMed=10198167;
RA   Osborne L.R., Campbell T., Daradich A., Scherer S.W., Tsui L.-C.;
RT   "Identification of a putative transcription factor gene (WBSCR11) that
RT   is commonly deleted in Williams-Beuren syndrome.";
RL   Genomics 57:279-284(1999).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Fetal brain;
RX   MEDLINE=20037629; PubMed=10573005;
RA   Tassabehji M., Carette M., Wilmot C., Donnai D., Read A.P.,
RA   Metcalfe K.;
RT   "A transcription factor involved in skeletal muscle gene expression is
RT   deleted in patients with Williams syndrome.";
RL   Eur. J. Hum. Genet. 7:737-747(1999).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RX   MEDLINE=20044629; PubMed=10575229;
RA   Franke Y., Peoples R.J., Francke U.;
RT   "Identification of GTF2IRD1, a putative transcription factor within
RT   the Williams-Beuren syndrome deletion at 7q11.23.";
RL   Cytogenet. Cell Genet. 86:296-304(1999).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM 1), INTERACTION WITH RB1, AND MUTAGENESIS.
RC   TISSUE=Cervical carcinoma, Fetal spleen, and Placenta;
RX   MEDLINE=20115113; PubMed=10642537;
RA   Yan X., Zhao X., Qian M., Guo N., Gong X., Zhu X.;
RT   "Characterization and gene structure of a novel retinoblastoma-
RT   protein-associated protein similar to the transcription regulator
RT   TFII-I.";
RL   Biochem. J. 345:749-757(2000).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Uterus;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [7]
RP   FUNCTION.
RX   MEDLINE=21332325; PubMed=11438732;
RA   Tussie-Luna M.I., Bayarsaihan D., Ruddle F.H., Roy A.L.;
RT   "Repression of TFII-I-dependent transcription by nuclear exclusion.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:7789-7794(2001).
CC   -!- FUNCTION: May be a transcription regulator involved in cell-cycle
CC       progression and skeletal muscle differentiation. May repress GTF2I
CC       transcriptional functions, by preventing its nuclear residency, or
CC       by inhibiting its transcriptional activation. May contribute to
CC       slow-twitch fiber type specificity during myogenesis and in
CC       regenerating muscles. Binds troponin I slow-muscle fiber enhancer
CC       (USE B1). Binds specifically and with high afinity to the EFG
CC       sequences derived from the early enhancer of HOXC8 (By
CC       similarity).
CC   -!- SUBUNIT: Interacts with the retinoblastoma protein (RB1) via its
CC       C-terminus.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q9UHL9-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q9UHL9-2; Sequence=VSP_003873;
CC   -!- TISSUE SPECIFICITY: Highly expressed in adult skeletal muscle,
CC       heart, fibroblast, bone and fetal tissues. Expressed at lower
CC       levels in all other tissues tested.
CC   -!- DEVELOPMENTAL STAGE: Highly expressed in developing and
CC       regenerating muscles, at the time of myofiber diversification.
CC   -!- DOMAIN: The N-terminal half may have an activating activity.
CC   -!- DISEASE: Haploinsufficiency of GTF2IRD1 may be the cause of
CC       certain cardiovascular and musculo-skeletal abnormalities observed
CC       in Williams-Beuren syndrome (WBS), a rare developmental disorder.
CC       It is a contiguous gene deletion syndrome involving genes from
CC       chromosome band 7q11.23.
CC   -!- SIMILARITY: BELONGS TO THE TFII-I FAMILY.
CC   -!- SIMILARITY: Contains 5 GTF2I-like repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF118270; AAD14687.2; -.
DR   EMBL; AF104923; AAD27668.1; -.
DR   EMBL; AF151354; AAF19786.1; -.
DR   EMBL; AF156489; AAF17358.1; -.
DR   EMBL; AF089107; AAF21796.1; -.
```

```
DR   EMBL; BC018136; AAH18136.1; -.
DR   Genew; HGNC:4661; GTF2IRD1.
DR   MIM; 604318; -.
DR   MIM; 194050; -.
DR   GO; GO:0005634; C:nucleus; NAS.
DR   GO; GO:0003705; F:RNA polymerase II transcription factor acti. . .; NAS.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR   InterPro; IPR004212; GTF2I.
DR   Pfam; PF02946; GTF2I; 5.
KW   Transcription regulation; Developmental protein; DNA-binding;
KW   Nuclear protein; Repeat; Alternative splicing; Polymorphism;
KW   Williams-Beuren syndrome.
FT   REPEAT      128    203       GTF2I 1.
FT   REPEAT      351    426       GTF2I 2.
FT   REPEAT      565    640       GTF2I 3.
FT   REPEAT      705    780       GTF2I 4.
FT   REPEAT      802    877       GTF2I 5.
FT   DOMAIN      898    905       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN      906    930       SER-RICH.
FT   VARSPLIC    656    670       Missing (in isoform 2).
FT                                /FTId=VSP_003873.
FT   VARIANT     652    652       M -> V (IN dbSNP:2301895).
FT                                /FTId=VAR_013446.
FT   MUTAGEN     898    959       MISSING: CYTOPLASMIC LOCALIZATION.
FT   CONFLICT    111    111       G -> S (IN REF. 1 AND 2).
FT   CONFLICT    378    378       R -> Q (IN REF. 5).
SQ   SEQUENCE   959 AA;  106057 MW;  7DA3097879701540 CRC64;

  Query Match             41.3%;  Score 43;  DB 1;  Length 959;
  Best Local Similarity   50.0%;  Pred. No. 75;
  Matches    7;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDFGSLAP 14
              || |  |: | | |
Db        480 CGPGTSGELGGLRP 493


RESULT 15
FLIH_TREPA
ID   FLIH_TREPA     STANDARD;      PRT;   309 AA.
AC   O83416;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Flagellar assembly protein fliH.
GN   FLIH OR TP0401.
OS   Treponema pallidum.
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX   NCBI_TaxID=160;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Nichols;
RX   MEDLINE=98332770; PubMed=9665876;
RA   Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA   Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA   Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA   Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA   McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA   Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA   Venter J.C.;
RT   "Complete genome sequence of Treponema pallidum, the syphilis
RT   spirochete.";
RL   Science 281:375-388(1998).
CC   -!- FUNCTION: NEEDED FOR FLAGELLAR REGROWTH AND ASSEMBLY (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: TO FLIH IN OTHER BACTERIA AND ALSO TO FLIJ.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001218; AAC65389.1; -.
DR   PIR; H71328; H71328.
DR   TIGR; TP0401; -.
DR   InterPro; IPR000563; Flag_FliH.
DR   Pfam; PF02108; FliH; 1.
KW   Flagella; Complete proteome.
SQ   SEQUENCE   309 AA;  35339 MW;  213978124489C81F CRC64;

  Query Match             40.4%;  Score 42;  DB 1;  Length 309;
  Best Local Similarity   52.6%;  Pred. No. 35;
  Matches   10;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              |  |  ||| :  ||| ||
Db        267 GCVVETDFGEIDARVASQL 285


Search completed: October  9, 2003, 10:14:57
Job time : 12.3333 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:09:40 ; Search time 49.3333 Seconds
(without alignments)
104.616 Million cell updates/sec

Title: US-09-549-186B-7
Perfect score: 104
Sequence: 1 CGSGVRGDFGSLAPRVARQL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	71	68.3	213	12	Q912L2	Q912l2 foot-and-mo
2	71	68.3	213	12	Q67446	Q67446 foot-and-mo
3	71	68.3	213	12	Q67448	Q67448 foot-and-mo
4	71	68.3	213	12	Q65093	Q65093 foot-and-mo
5	70	67.3	20	12	Q9PXE4	Q9pxe4 foot-and-mo
6	68	65.4	213	12	Q98VZ6	Q98vz6 foot-and-mo
7	67	64.4	210	12	Q67438	Q67438 foot-and-mo
8	67	64.4	969	12	Q9Q2N6	Q9q2n6 foot-and-mo
9	66	63.5	169	12	Q67442	Q67442 foot-and-mo
10	65	62.5	126	12	Q8V443	Q8v443 foot-and-mo
11	62	59.6	143	12	Q9YPU0	Q9ypu0 foot-and-mo
12	62	59.6	208	12	Q8JUR8	Q8jur8 foot-and-mo
13	62	59.6	211	12	Q8JUP0	Q8jup0 foot-and-mo
14	62	59.6	212	12	Q8JUP1	Q8jup1 foot-and-mo
15	62	59.6	213	12	Q8JUS6	Q8jus6 foot-and-mo
16	62	59.6	213	12	Q8JUM1	Q8jum1 foot-and-mo
17	62	59.6	213	12	Q8JUQ9	Q8juq9 foot-and-mo
18	62	59.6	213	12	Q8JUP7	Q8jup7 foot-and-mo
19	62	59.6	213	12	Q8JUP4	Q8jup4 foot-and-mo
20	62	59.6	213	12	Q912L1	Q912l1 foot-and-mo
21	62	59.6	213	12	Q8JUQ5	Q8juq5 foot-and-mo
22	62	59.6	213	12	Q8JUL1	Q8jul1 foot-and-mo
23	62	59.6	213	12	Q8JUN1	Q8jun1 foot-and-mo
24	62	59.6	213	12	Q8JUP5	Q8jup5 foot-and-mo
25	62	59.6	213	12	Q8JUN0	Q8jun0 foot-and-mo
26	62	59.6	213	12	Q8JUP3	Q8jup3 foot-and-mo
27	62	59.6	213	12	Q8JUQ2	Q8juq2 foot-and-mo
28	62	59.6	213	12	Q8JUQ1	Q8juq1 foot-and-mo
29	62	59.6	213	12	Q8JUL0	Q8jul0 foot-and-mo
30	62	59.6	213	12	Q8JUS7	Q8jus7 foot-and-mo
31	62	59.6	213	12	Q8JUM0	Q8jum0 foot-and-mo
32	62	59.6	213	12	Q8JUM3	Q8jum3 foot-and-mo
33	62	59.6	213	12	Q8JUQ8	Q8juq8 foot-and-mo
34	62	59.6	213	12	Q8JUN9	Q8jun9 foot-and-mo
35	62	59.6	213	12	Q8JUR0	Q8jur0 foot-and-mo
36	62	59.6	213	12	Q8JUS5	Q8jus5 foot-and-mo
37	60	57.7	197	12	Q8JUT2	Q8jut2 foot-and-mo
38	60	57.7	213	12	Q8JUQ4	Q8juq4 foot-and-mo
39	60	57.7	652	12	Q9Q2N8	Q9q2n8 foot-and-mo
40	59	56.7	125	12	Q8V442	Q8v442 foot-and-mo
41	59	56.7	213	12	Q8JUM9	Q8jum9 foot-and-mo
42	59	56.7	213	12	Q912K6	Q912k6 foot-and-mo
43	59	56.7	213	12	Q66924	Q66924 foot-and-mo
44	59	56.7	213	12	Q912K4	Q912k4 foot-and-mo
45	59	56.7	213	12	Q912L0	Q91210 foot-and-mo

ALIGNMENTS

RESULT 1
Q912L2
ID Q912L2 PRELIMINARY; PRT; 213 AA.
AC Q912L2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE VP1 protein (Fragment).
GN 1D.
OS Foot-and-mouth disease virus A.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Aphthovirus.
OX NCBI_TaxID=12111;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A Arg/68;
RX MEDLINE=21580796; PubMed=11724271;
RA Konig G.A., Blanco C., Knowles N.J., Palma E.L., Maradei E.,
RA Piccone M.E.;
RT "Phylogenetic analysis of foot-and-mouth disease viruses isolated in
RT Argentina.";
RL Virus Genes 23:175-182(2001).
DR EMBL; AJ308694; CAC48168.1; -.
DR InterPro; IPR004080; FMDVP1coat.
DR InterPro; IPR001676; Rhv.
DR Pfam; PF00073; rhv; 1.
DR PRINTS; PR01542; FMDVP1COAT.
FT NON_TER 1 1
FT NON_TER 213 213
SQ SEQUENCE 213 AA; 23488 MW; 203EFCB8AB45EECE CRC64;

Query Match 68.3%; Score 71; DB 12; Length 213;
Best Local Similarity 78.9%; Pred. No. 0.0023;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GSGVRGDFGSLAPRVARQL 20
||| ||| |||| |||:||
Db 140 GSGRRGDMGSLAARVAKQL 158

```
RESULT 2
Q67446
ID   Q67446      PRELIMINARY;      PRT;   213 AA.
AC   Q67446; Q67447;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Foot and mouth disease virus (Serotype Aven76), capsid protein VP1
DE   mRNA (Fragment).
OS   Foot-and-mouth disease virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12110;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85190547; PubMed=2986125;
RA   Weddell G.N., Yansura D.G., Dowbenko D.J., Hoatlin M.E., Grubman M.J.,
RA   Moore D.M., Kleid D.G.;
RT   "Sequence variation in the gene for the immunogenic capsid protein VP1
RT   of foot-and-mouth disease virus type A.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:2618-2622(1985).
DR   EMBL; K03344; AAA42600.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     213    213
SQ   SEQUENCE   213 AA;  23311 MW;  C4ED4E0116A0DB8A CRC64;

  Query Match             68.3%;  Score 71;  DB 12;  Length 213;
  Best Local Similarity   78.9%;  Pred. No. 0.0023;
  Matches   15;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||| ||| |||| |||:||
Db        140 GSGRRGDMGSLAARVAKQL 158


RESULT 3
Q67448
ID   Q67448      PRELIMINARY;      PRT;   213 AA.
AC   Q67448; Q67449;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Foot and mouth disease virus (Serotype Aarg79), capsid protein VP1
DE   mRNA (Fragment).
OS   Foot-and-mouth disease virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12110;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85190547; PubMed=2986125;
RA   Weddell G.N., Yansura D.G., Dowbenko D.J., Hoatlin M.E., Grubman M.J.,
RA   Moore D.M., Kleid D.G.;
RT   "Sequence variation in the gene for the immunogenic capsid protein VP1
RT   of foot-and-mouth disease virus type A.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:2618-2622(1985).
DR   EMBL; K03345; AAA42601.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     213    213
SQ   SEQUENCE   213 AA;  23345 MW;  546C7FF78CD45CC0 CRC64;

  Query Match             68.3%;  Score 71;  DB 12;  Length 213;
  Best Local Similarity   78.9%;  Pred. No. 0.0023;
  Matches   15;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||| ||| |||| |||:||
Db        140 GSGRRGDMGSLAARVAKQL 158


RESULT 4
Q65093
ID   Q65093      PRELIMINARY;      PRT;   213 AA.
AC   Q65093;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Foot and mouth disease virus strain A Venceslau VP1 (Fragment).
OS   Foot-and-mouth disease virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12110;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85077620; PubMed=6096217;
RA   Cheung A., Whitehead P., Weiss S., Kupper H.;
RT   "Nucleotide sequence of the VP1 gene of the foot-and-mouth disease
RT   virus strain A Venceslau.";
RL   Gene 30:241-245(1984).
DR   EMBL; M12905; AAA42663.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     213    213
SQ   SEQUENCE   213 AA;  23327 MW;  048088D9571D7073 CRC64;

  Query Match             68.3%;  Score 71;  DB 12;  Length 213;
  Best Local Similarity   78.9%;  Pred. No. 0.0023;
  Matches   15;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||| ||| |||| |||:||
Db        140 GSGRRGDMGSLAARVAKQL 158


RESULT 5
Q9PXE4
ID   Q9PXE4      PRELIMINARY;      PRT;    20 AA.
AC   Q9PXE4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   VP1 protein (Fragment).
OS   Foot-and-mouth disease virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12110;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96100820; PubMed=7483796;
RA   Piatti P., Hassard S., Newman J.F., Brown F.;
RT   "Antigenic variants in a plaque-isolate of foot-and-mouth disease
RT   virus: implications for vaccine production.";
RL   Vaccine 13:781-784(1995).
SQ   SEQUENCE   20 AA;  1925 MW;  E91F87C99C19D7DB CRC64;

  Query Match             67.3%;  Score 70;  DB 12;  Length 20;
  Best Local Similarity   84.2%;  Pred. No. 0.00026;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||||| |||| |||| |
Db          1 GSGVRGDSGSLALRVARGL 19
```

```
RESULT 6
Q98VZ6
ID   Q98VZ6      PRELIMINARY;      PRT;   213 AA.
AC   Q98VZ6;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   VP1 (Fragment).
GN   1D.
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A76/Argentina/76;
RX   MEDLINE=21580796; PubMed=11724271;
RA   Konig G.A., Blanco C., Knowles N.J., Palma E.L., Maradei E.,
RA   Piccone M.E.;
RT   "Phylogenetic analysis of foot-and-mouth disease viruses isolated in
RT   Argentina.";
RL   Virus Genes 23:175-182(2001).
DR   EMBL; AJ409219; CAC34727.1; -.
DR   HSSP; Q88571; 1TME.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     213    213
SQ   SEQUENCE   213 AA;  23494 MW;  92CBE4DBC885F3C9 CRC64;

  Query Match             65.4%;  Score 68;  DB 12;  Length 213;
  Best Local Similarity   78.9%;  Pred. No. 0.0071;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||  ||| |||| ||||||
Db        140 GSARRGDMGSLAARVARQL 158


RESULT 7
Q67438
ID   Q67438      PRELIMINARY;      PRT;   210 AA.
AC   Q67438; Q67439;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Foot and mouth disease virus (Serotype A24), capsid protein VP1 mRNA
DE   (Fragment).
OS   Foot-and-mouth disease virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12110;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85190547; PubMed=2986125;
RA   Weddell G.N., Yansura D.G., Dowbenko D.J., Hoatlin M.E., Grubman M.J.,
RA   Moore D.M., Kleid D.G.;
RT   "Sequence variation in the gene for the immunogenic capsid protein VP1
RT   of foot-and-mouth disease virus type A.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:2618-2622(1985).
DR   EMBL; K03340; AAA42595.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     210    210
SQ   SEQUENCE   210 AA;  23089 MW;  1C96CA1C5E4F6A60 CRC64;

  Query Match             64.4%;  Score 67;  DB 12;  Length 210;
  Best Local Similarity   73.7%;  Pred. No. 0.01;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||| ||| |||| || :||
Db        137 GSGRRGDMGSLAARVVKQL 155


RESULT 8
Q9Q2N6
ID   Q9Q2N6      PRELIMINARY;      PRT;   969 AA.
AC   Q9Q2N6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Polyprotein (Fragment).
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A24/Cruzeiro/Brazil/55;
RA   Abrams C.C.;
RT   "Nucleotide sequence of foot-and-mouth disease virus
RT   A24/Cruzeiro/Brazil/55 from the poly(C) tract to 2B.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A24/Cruzeiro/Brazil/55;
RA   Knowles N.J., Samuel A.R., Aktas S., Rowe C.A., Abrams C.C.,
RA   Newman J.W.I., King A.M.Q.;
RT   "Phylogenetic comparison of the capsid-coding region of all seven
RT   foot-and-mouth disease virus serotypes.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ251476; CAB62583.1; -.
DR   HSSP; Q88571; 1TME.
DR   MEROPS; C28.001; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 3.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   CHAIN         1   >201       LEADER.
FT   CHAIN       202   >286       1A (VP4).
FT   CHAIN       287   >504       1B (VP2).
FT   CHAIN       505   >725       1C (VP3).
FT   CHAIN       726   >938       1D (VP1).
FT   CHAIN       939   >954       2A.
FT   CHAIN       955   >969       2B.
FT   NON_TER     969    969
SQ   SEQUENCE   969 AA;  107629 MW;  7DD5D908FCFF89FD CRC64;

  Query Match             64.4%;  Score 67;  DB 12;  Length 969;
  Best Local Similarity   73.7%;  Pred. No. 0.053;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||| ||| |||| || :||
Db        865 GSGRRGDMGSLAARVVKQL 883


RESULT 9
Q67442
ID   Q67442      PRELIMINARY;      PRT;   169 AA.
AC   Q67442; Q67443;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Foot and mouth disease virus (Serotype A32), capsid protein VP1 mRNA
DE   (Fragment).
OS   Foot-and-mouth disease virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12110;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85190547; PubMed=2986125;
RA   Weddell G.N., Yansura D.G., Dowbenko D.J., Hoatlin M.E., Grubman M.J.,
RA   Moore D.M., Kleid D.G.;
RT   "Sequence variation in the gene for the immunogenic capsid protein VP1
RT   of foot-and-mouth disease virus type A.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:2618-2622(1985).
DR   EMBL; K03342; AAA42598.1; -.
DR   HSSP; Q88571; 1TME.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 2.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     169    169
SQ   SEQUENCE   169 AA;  18499 MW;  9E8AEDC94AC57F67 CRC64;

  Query Match             63.5%;  Score 66;  DB 12;  Length 169;
  Best Local Similarity   73.7%;  Pred. No. 0.012;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||| ||| | || |||:||
Db         96 GSGRRGDLGPLAARVAKQL 114


RESULT 10
Q8V443
ID   Q8V443      PRELIMINARY;      PRT;   126 AA.
AC   Q8V443;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   VP1 (Fragment).
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KEN/1/76;
RA   van Rensburg H.G., Bastos A.D., Haydon D., Joubert F., Nel L.H.;
RT   "Gene heterogeneity in the leader and 3C proteinases of foot-and-mouth
RT   disease virus.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF390861; AAL73359.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     126    126
SQ   SEQUENCE   126 AA;  13693 MW;  82675FAC45F713BB CRC64;

  Query Match             62.5%;  Score 65;  DB 12;  Length 126;
  Best Local Similarity   77.8%;  Pred. No. 0.012;
  Matches   14;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 SGVRGDFGSLAPRVARQL 20
              || ||| |:|| ||||||
Db         54 SGRRGDMGALAARVARQL 71


RESULT 11
Q9YPU0
ID   Q9YPU0      PRELIMINARY;      PRT;   143 AA.
AC   Q9YPU0;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   VP1 protein (Fragment).
GN   VP1.
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NM/XZ/64;
RA   Liu Z., Zhao Q., Liu W., Xie Q.;
RT   "Analysis of VP1 coding nucleotide sequences of six strains of foot-
RT   and-mouth disease virus type A.";
RL   Ping Tu Hsueh Pao 14:60-67(1998).
DR   EMBL; AJ131664; CAA10474.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     143    143
SQ   SEQUENCE   143 AA;  15881 MW;  CF0CC1858CC331E0 CRC64;

  Query Match             59.6%;  Score 62;  DB 12;  Length 143;
  Best Local Similarity   72.2%;  Pred. No. 0.042;
  Matches   13;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 SGVRGDFGSLAPRVARQL 20
              :| ||| |||| |||:||
Db         71 AGRRGDLGSLAARVAKQL 88


RESULT 12
Q8JUR8
ID   Q8JUR8      PRELIMINARY;      PRT;   208 AA.
AC   Q8JUR8;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   1D protein (Fragment).
GN   POL.
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/IND/170/88;
RX   MEDLINE=21954626; PubMed=11958451;
RA   Tosh C., Sanyal A., Hemadri D., Venkataramanan R.;
RT   "Phylogenetic analysis of serotype A foot-and-mouth disease virus
RT   isolated in India between 1977 and 2000.";
RL   Arch. Virol. 147:493-513(2002).
DR   EMBL; AF390612; AAM63984.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  22607 MW;  6CE616F86E60CF7E CRC64;

  Query Match             59.6%;  Score 62;  DB 12;  Length 208;
  Best Local Similarity   68.4%;  Pred. No. 0.064;
  Matches   13;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              |:| ||| |:|| ||| ||
Db        140 GAGRRGDLGALAARVAAQL 158


RESULT 13
Q8JUP0
ID   Q8JUP0      PRELIMINARY;      PRT;   211 AA.
AC   Q8JUP0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
```

```
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   1D protein (Fragment).
GN   POL.
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/IND/302/88;
RX   MEDLINE=21954626; PubMed=11958451;
RA   Tosh C., Sanyal A., Hemadri D., Venkataramanan R.;
RT   "Phylogenetic analysis of serotype A foot-and-mouth disease virus
RT   isolated in India between 1977 and 2000.";
RL   Arch. Virol. 147:493-513(2002).
DR   EMBL; AF390641; AAM64012.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     211    211
SQ   SEQUENCE   211 AA;  23100 MW;  CFBFC94CCD9E2FE3 CRC64;

  Query Match             59.6%;  Score 62;  DB 12;  Length 211;
  Best Local Similarity   68.4%;  Pred. No. 0.065;
  Matches   13;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              |:| ||| |:|| ||| ||
Db        140 GTGRRGDLGALAARVAAQL 158


RESULT 14
Q8JUP1
ID   Q8JUP1      PRELIMINARY;      PRT;   212 AA.
AC   Q8JUP1;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   1D protein (Fragment).
GN   POL.
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/IND/299/99;
RX   MEDLINE=21954626; PubMed=11958451;
RA   Tosh C., Sanyal A., Hemadri D., Venkataramanan R.;
RT   "Phylogenetic analysis of serotype A foot-and-mouth disease virus
RT   isolated in India between 1977 and 2000.";
RL   Arch. Virol. 147:493-513(2002).
DR   EMBL; AF390640; AAM64011.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     212    212
SQ   SEQUENCE   212 AA;  23150 MW;  E6B690890042CD3A CRC64;

  Query Match             59.6%;  Score 62;  DB 12;  Length 212;
  Best Local Similarity   68.4%;  Pred. No. 0.065;
  Matches   13;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              |:| ||| |:|| ||| ||
Db        139 GTGRRGDLGALAARVAAQL 157


RESULT 15
Q8JUS6
ID   Q8JUS6      PRELIMINARY;      PRT;   213 AA.
AC   Q8JUS6;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   1D protein (Fragment).
GN   POL.
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/IND/138/99;
RX   MEDLINE=21954626; PubMed=11958451;
RA   Tosh C., Sanyal A., Hemadri D., Venkataramanan R.;
RT   "Phylogenetic analysis of serotype A foot-and-mouth disease virus
RT   isolated in India between 1977 and 2000.";
RL   Arch. Virol. 147:493-513(2002).
DR   EMBL; AF390604; AAM63976.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     213    213
SQ   SEQUENCE   213 AA;  23106 MW;  2C7F6C9ED1039C0B CRC64;

  Query Match             59.6%;  Score 62;  DB 12;  Length 213;
  Best Local Similarity   68.4%;  Pred. No. 0.065;
  Matches   13;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              |:| ||| |:|| ||| ||
Db        140 GTGRRGDLGALAARVAAQL 158


Search completed: October  9, 2003, 10:17:34
Job time : 52.3333 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:11:24 ; Search time 17.3333 Seconds
(without alignments)
48.820 Million cell updates/sec

Title: US-09-549-186B-7
Perfect score: 104
Sequence: 1 CGSGVRGDFGSLAPRVARQL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	104	100.0	20	4	US-08-716-249-7	Sequence 7, Appli
2	95	91.3	35	3	US-09-100-600A-1	Sequence 1, Appli
3	95	91.3	46	3	US-09-100-600A-38	Sequence 38, Appl
4	95	91.3	75	3	US-09-100-600A-28	Sequence 28, Appl
5	94	90.4	20	4	US-08-716-249-8	Sequence 8, Appli
6	92	88.5	31	1	US-08-418-716A-4	Sequence 4, Appli
7	91	87.5	31	1	US-08-418-716A-5	Sequence 5, Appli
8	91	87.5	31	1	US-08-418-716A-6	Sequence 6, Appli
9	89.5	86.1	19	4	US-08-716-249-2	Sequence 2, Appli
10	89	85.6	31	1	US-08-418-716A-2	Sequence 2, Appli
11	88	84.6	31	1	US-08-418-716A-1	Sequence 1, Appli
12	88	84.6	31	1	US-08-418-716A-7	Sequence 7, Appli
13	87	83.7	35	3	US-09-100-600A-2	Sequence 2, Appli
14	87	83.7	46	3	US-09-100-600A-3	Sequence 3, Appli
15	87	83.7	65	3	US-09-100-600A-27	Sequence 27, Appl
16	87	83.7	75	3	US-09-100-600A-29	Sequence 29, Appl
17	86	82.7	20	4	US-08-716-249-9	Sequence 9, Appli
18	76	73.1	31	1	US-08-418-716A-3	Sequence 3, Appli
19	71.5	68.8	19	4	US-08-716-249-1	Sequence 1, Appli
20	71	68.3	35	3	US-09-100-600A-8	Sequence 8, Appli
21	71	68.3	46	3	US-09-100-600A-42	Sequence 42, Appl
22	71	68.3	46	3	US-09-100-600A-49	Sequence 49, Appl
23	71	68.3	46	3	US-09-100-600A-50	Sequence 50, Appl
24	67	64.4	46	3	US-09-100-600A-43	Sequence 43, Appl
25	67	64.4	46	3	US-09-100-600A-44	Sequence 44, Appl
26	67	64.4	46	3	US-09-100-600A-48	Sequence 48, Appl
27	66	63.5	46	3	US-09-100-600A-45	Sequence 45, Appl
28	63	60.6	35	3	US-09-100-600A-9	Sequence 9, Appli
29	63	60.6	55	3	US-09-100-600A-33	Sequence 33, Appl
30	60	57.7	46	3	US-09-100-600A-39	Sequence 39, Appl
31	59	56.7	46	3	US-09-100-600A-46	Sequence 46, Appl
32	55	52.9	45	3	US-09-100-600A-55	Sequence 55, Appl
33	54	51.9	46	3	US-09-100-600A-47	Sequence 47, Appl
34	53	51.0	45	3	US-09-100-600A-60	Sequence 60, Appl
35	50	48.1	45	3	US-09-100-600A-51	Sequence 51, Appl
36	50	48.1	45	3	US-09-100-600A-62	Sequence 62, Appl
37	50	48.1	45	3	US-09-100-600A-64	Sequence 64, Appl
38	50	48.1	45	3	US-09-100-600A-65	Sequence 65, Appl
39	50	48.1	46	3	US-09-100-600A-40	Sequence 40, Appl
40	50	48.1	46	3	US-09-100-600A-41	Sequence 41, Appl
41	49	47.1	387	4	US-09-252-991A-17881	Sequence 17881, A
42	48	46.2	45	3	US-09-100-600A-56	Sequence 56, Appl
43	48	46.2	387	4	US-09-252-991A-18477	Sequence 18477, A
44	46	44.2	20	6	5204096-32	Patent No. 5204096
45	46	44.2	45	3	US-09-100-600A-52	Sequence 52, Appl

ALIGNMENTS

RESULT 1
US-08-716-249-7
; Sequence 7, Application US/08716249
; Patent No. 6455244
; GENERAL INFORMATION:
; APPLICANT: Guichard, Gilles
; APPLICANT: Muller, Sylviane
; APPLICANT: Briand, Jean-Paul
; APPLICANT: Regenmortel, Marc
; TITLE OF INVENTION: Retropeptides, Antibodies Thereto, and
; TITLE OF INVENTION: Uses Thereof for Vaccination and In Vitro Diagnosis
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spencer & Frank
; STREET: 1100 New York Avenue, Suite 300E
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/716,249
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/00292
; FILING DATE: 13-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Calvetti, Frederick F.
; REGISTRATION NUMBER: 28,557
; REFERENCE/DOCKET NUMBER: GROFO 7001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)414-4000
; TELEFAX: (202)414-4040
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-716-249-7

Query Match 100.0%; Score 104; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CGSGVRGDFGSLAPRVARQL 20
              ||||||||||||||||||||
Db          1 CGSGVRGDFGSLAPRVARQL 20


RESULT 2
US-09-100-600A-1
; Sequence 1, Application US/09100600A
; Patent No. 6107021
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Shen, Ming
;    TITLE OF INVENTION:  SYNTHETIC PEPTIDE VACCINES FOR
;    TITLE OF INVENTION:  FOOT-AND-MOUTH DISEASE
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,600A
;      FILING DATE:  20-Jun-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria C.H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4156
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  35 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-600A-1

  Query Match             91.3%;  Score 95;  DB 3;  Length 35;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-08;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||||||||||||||||||
Db          7 GSGVRGDFGSLAPRVARQL 25


RESULT 3
US-09-100-600A-38
; Sequence 38, Application US/09100600A
; Patent No. 6107021
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Shen, Ming
;    TITLE OF INVENTION:  SYNTHETIC PEPTIDE VACCINES FOR
;    TITLE OF INVENTION:  FOOT-AND-MOUTH DISEASE
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,600A
;      FILING DATE:  20-Jun-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria C.H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4156
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-600A-38

  Query Match             91.3%;  Score 95;  DB 3;  Length 46;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-08;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||||||||||||||||||
Db         16 GSGVRGDFGSLAPRVARQL 34


RESULT 4
US-09-100-600A-28
; Sequence 28, Application US/09100600A
; Patent No. 6107021
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Shen, Ming
;    TITLE OF INVENTION:  SYNTHETIC PEPTIDE VACCINES FOR
;    TITLE OF INVENTION:  FOOT-AND-MOUTH DISEASE
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,600A
;      FILING DATE:  20-Jun-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria C.H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4156
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  75 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-600A-28
```

```
  Query Match             91.3%;  Score 95;  DB 3;  Length 75;
  Best Local Similarity   100.0%;  Pred. No. 9.6e-08;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||||||||||||||||||
Db         47 GSGVRGDFGSLAPRVARQL 65


RESULT 5
US-08-716-249-8
; Sequence 8, Application US/08716249
; Patent No. 6455244
;  GENERAL INFORMATION:
;    APPLICANT:  Guichard, Gilles
;    APPLICANT:  Muller, Sylviane
;    APPLICANT:  Briand, Jean-Paul
;    APPLICANT:  Regenmortel, Marc
;    TITLE OF INVENTION:  Retropeptides, Antibodies Thereto, and
;    TITLE OF INVENTION:  Uses Thereof for Vaccination and In Vitro Diagnosis
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spencer & Frank
;      STREET:  1100 New York Avenue, Suite 300E
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/716,249
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/00292
;      FILING DATE:  13-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Calvetti, Frederick F.
;      REGISTRATION NUMBER:  28,557
;      REFERENCE/DOCKET NUMBER:  GROFO 7001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)414-4000
;      TELEFAX:  (202)414-4040
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  YES
US-08-716-249-8

  Query Match             90.4%;  Score 94;  DB 4;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 3.4e-08;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDFGSLAPRVARQL 20
              |||||||||||||  ||||||
Db          1 CGSGVRGDFGSLALRVARQL 20


RESULT 6
US-08-418-716A-4
; Sequence 4, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;    APPLICANT:  Mason, Peter W.
;    APPLICANT:  Baxt, Barry
;    APPLICANT:  Reider, Elizabeth
;    APPLICANT:  Berinstein, Analia
;    APPLICANT:  Kang, Angray S
;    TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;    TITLE OF INVENTION:  Viruses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janelle S. Graeter
;      STREET:  Rm. 411, Bldg. 005, BARC-W
;      CITY:  Beltsville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20705
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/418,716A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Graeter, Janelle S.
;      REGISTRATION NUMBER:  35,024
;      REFERENCE/DOCKET NUMBER:  0137.94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 504-6629
;      TELEFAX:  (301) 504-5060
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:
;      ORGANISM:  Foot and mouth disease virus
US-08-418-716A-4

  Query Match             88.5%;  Score 92;  DB 1;  Length 31;
  Best Local Similarity   94.7%;  Pred. No. 1.1e-07;
  Matches   18;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||:|||||||||||||||
Db         11 GSGVKGDFGSLAPRVARQL 29


RESULT 7
US-08-418-716A-5
; Sequence 5, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;    APPLICANT:  Mason, Peter W.
;    APPLICANT:  Baxt, Barry
;    APPLICANT:  Reider, Elizabeth
;    APPLICANT:  Berinstein, Analia
;    APPLICANT:  Kang, Angray S
;    TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;    TITLE OF INVENTION:  Viruses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janelle S. Graeter
;      STREET:  Rm. 411, Bldg. 005, BARC-W
;      CITY:  Beltsville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20705
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/418,716A
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Graeter, Janelle S.
;       REGISTRATION NUMBER:  35,024
;       REFERENCE/DOCKET NUMBER:  0137.94
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 504-6629
;       TELEFAX:  (301) 504-5060
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  31 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     HYPOTHETICAL:  NO
;     FRAGMENT TYPE:  internal
;     ORIGINAL SOURCE:
;       ORGANISM:  Foot and mouth disease virus
US-08-418-716A-5

  Query Match             87.5%;  Score 91;  DB 1;  Length 31;
  Best Local Similarity   94.7%;  Pred. No. 1.6e-07;
  Matches   18;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||||:||||||||||||
Db         11 GSGVRGEFGSLAPRVARQL 29


RESULT 8
US-08-418-716A-6
; Sequence 6, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;    APPLICANT:  Mason, Peter W.
;    APPLICANT:  Baxt, Barry
;    APPLICANT:  Reider, Elizabeth
;    APPLICANT:  Berinstein, Analia
;    APPLICANT:  Kang, Angray S
;    TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;    TITLE OF INVENTION:  Viruses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janelle S. Graeter
;      STREET:  Rm. 411, Bldg. 005, BARC-W
;      CITY:  Beltsville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20705
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/418,716A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Graeter, Janelle S.
;      REGISTRATION NUMBER:  35,024
;      REFERENCE/DOCKET NUMBER:  0137.94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 504-6629
;      TELEFAX:  (301) 504-5060
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:
;      ORGANISM:  Foot and mouth disease virus
US-08-418-716A-6

  Query Match             87.5%;  Score 91;  DB 1;  Length 31;
  Best Local Similarity   94.7%;  Pred. No. 1.6e-07;
  Matches   18;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||||:||||||||||||
Db         11 GSGVRGEFGSLAPRVARQL 29


RESULT 9
US-08-716-249-2
; Sequence 2, Application US/08716249
; Patent No. 6455244
;  GENERAL INFORMATION:
;    APPLICANT:  Guichard, Gilles
;    APPLICANT:  Muller, Sylviane
;    APPLICANT:  Briand, Jean-Paul
;    APPLICANT:  Regenmortel, Marc
;    TITLE OF INVENTION:  Retropeptides, Antibodies Thereto, and
;    TITLE OF INVENTION:  Uses Thereof for Vaccination and In Vitro Diagnosis
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spencer & Frank
;      STREET:  1100 New York Avenue, Suite 300E
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/716,249
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/00292
;      FILING DATE:  13-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Calvetti, Frederick F.
;      REGISTRATION NUMBER:  28,557
;      REFERENCE/DOCKET NUMBER:  GROFO 7001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)414-4000
;      TELEFAX:  (202)414-4040
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-716-249-2

  Query Match             86.1%;  Score 89.5;  DB 4;  Length 19;
  Best Local Similarity   95.0%;  Pred. No. 1.6e-07;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 CGSGVRGDFGSLAPRVARQL 20
```

```
                    |||||||||||| ||||||||
Db          1 CGSGVRGDFGS-APRVARQL 19


RESULT 10
US-08-418-716A-2
; Sequence 2, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;    APPLICANT:  Mason, Peter W.
;    APPLICANT:  Baxt, Barry
;    APPLICANT:  Reider, Elizabeth
;    APPLICANT:  Berinstein, Analia
;    APPLICANT:  Kang, Angray S
;    TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;    TITLE OF INVENTION:  Viruses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janelle S. Graeter
;      STREET:  Rm. 411, Bldg. 005, BARC-W
;      CITY:  Beltsville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20705
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/418,716A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Graeter, Janelle S.
;      REGISTRATION NUMBER:  35,024
;      REFERENCE/DOCKET NUMBER:  0137.94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 504-6629
;      TELEFAX:  (301) 504-5060
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:
;      ORGANISM:  Foot and mouth disease virus
US-08-418-716A-2

  Query Match             85.6%;  Score 89;  DB 1;  Length 31;
  Best Local Similarity   94.7%;  Pred. No. 3.2e-07;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||| ||||||||||||||||
Db         11 GSGPRGDFGSLAPRVARQL 29


RESULT 11
US-08-418-716A-1
; Sequence 1, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;    APPLICANT:  Mason, Peter W.
;    APPLICANT:  Baxt, Barry
;    APPLICANT:  Reider, Elizabeth
;    APPLICANT:  Berinstein, Analia
;    APPLICANT:  Kang, Angray S
;    TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;    TITLE OF INVENTION:  Viruses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janelle S. Graeter
;      STREET:  Rm. 411, Bldg. 005, BARC-W
;      CITY:  Beltsville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20705
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/418,716A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Graeter, Janelle S.
;      REGISTRATION NUMBER:  35,024
;      REFERENCE/DOCKET NUMBER:  0137.94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 504-6629
;      TELEFAX:  (301) 504-5060
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:
;      ORGANISM:  Foot and mouth disease virus
US-08-418-716A-1

  Query Match             84.6%;  Score 88;  DB 1;  Length 31;
  Best Local Similarity   94.7%;  Pred. No. 4.6e-07;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||| ||||||||||||||||
Db         11 GSGDRGDFGSLAPRVARQL 29


RESULT 12
US-08-418-716A-7
; Sequence 7, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;    APPLICANT:  Mason, Peter W.
;    APPLICANT:  Baxt, Barry
;    APPLICANT:  Reider, Elizabeth
;    APPLICANT:  Berinstein, Analia
;    APPLICANT:  Kang, Angray S
;    TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;    TITLE OF INVENTION:  Viruses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janelle S. Graeter
;      STREET:  Rm. 411, Bldg. 005, BARC-W
;      CITY:  Beltsville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20705
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  US/08/418,716A
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Graeter, Janelle S.
;       REGISTRATION NUMBER:  35,024
;       REFERENCE/DOCKET NUMBER:  0137.94
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 504-6629
;       TELEFAX:  (301) 504-5060
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  31 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     HYPOTHETICAL:  NO
;     FRAGMENT TYPE:  internal
;     ORIGINAL SOURCE:
;       ORGANISM:  Foot and mouth disease virus
US-08-418-716A-7

  Query Match             84.6%;  Score 88;  DB 1;  Length 31;
  Best Local Similarity   89.5%;  Pred. No. 4.6e-07;
  Matches   17;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              ||||:|:|||||||||||||
Db         11 GSGVRGEFGSLAPRVARQL 29


RESULT 13
US-09-100-600A-2
; Sequence 2, Application US/09100600A
; Patent No. 6107021
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Shen, Ming
;    TITLE OF INVENTION:  SYNTHETIC PEPTIDE VACCINES FOR
;    TITLE OF INVENTION:  FOOT-AND-MOUTH DISEASE
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,600A
;      FILING DATE:  20-Jun-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria C.H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4156
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  35 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-600A-2

  Query Match             83.7%;  Score 87;  DB 3;  Length 35;
  Best Local Similarity   94.7%;  Pred. No. 7.4e-07;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              |||||||||||||||||| |
Db          7 GSGVRGDFGSLAPRVARCL 25


RESULT 14
US-09-100-600A-3
; Sequence 3, Application US/09100600A
; Patent No. 6107021
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Shen, Ming
;    TITLE OF INVENTION:  SYNTHETIC PEPTIDE VACCINES FOR
;    TITLE OF INVENTION:  FOOT-AND-MOUTH DISEASE
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,600A
;      FILING DATE:  20-Jun-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria C.H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4156
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-600A-3

  Query Match             83.7%;  Score 87;  DB 3;  Length 46;
  Best Local Similarity   94.7%;  Pred. No. 9.9e-07;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              |||||||||||||||||| |
Db         16 GSGVRGDFGSLAPRVARCL 34


RESULT 15
US-09-100-600A-27
; Sequence 27, Application US/09100600A
; Patent No. 6107021
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Shen, Ming
;    TITLE OF INVENTION:  SYNTHETIC PEPTIDE VACCINES FOR
;    TITLE OF INVENTION:  FOOT-AND-MOUTH DISEASE
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
```

```
;       STATE:  NY
;       COUNTRY:  USA
;       ZIP:  10154-0054
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  Windows 95
;       SOFTWARE:  Word 97
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/100,600A
;       FILING DATE:  20-Jun-1998
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Maria C.H. Lin
;       REGISTRATION NUMBER:  29,323
;       REFERENCE/DOCKET NUMBER:  1151-4156
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  212-758-4800
;       TELEFAX:  212-751-6849
;   INFORMATION FOR SEQ ID NO:  27:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  65 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-09-100-600A-27

  Query Match             83.7%;  Score 87;  DB 3;  Length 65;
  Best Local Similarity   94.7%;  Pred. No. 1.4e-06;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              |||||||||||||||||| |
Db         47 GSGVRGDFGSLAPRVARCL 65


Search completed: October  9, 2003, 10:19:37
Job time : 17.3333 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:14:31 ; Search time 21.3333 Seconds
(without alignments)
151.058 Million cell updates/sec

Title: US-09-549-186B-7
Perfect score: 104
Sequence: 1 CGSGVRGDFGSLAPRVARQL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	44	42.3	240	15	US-10-156-761-14172	Sequence 14172, A
2	44	42.3	920	10	US-09-983-531A-4	Sequence 4, Appli
3	44	42.3	998	15	US-10-101-464A-931	Sequence 931, App
4	43	41.3	312	15	US-10-207-655-186	Sequence 186, App
5	43	41.3	476	9	US-09-291-299A-3	Sequence 3, Appli
6	43	41.3	627	15	US-10-222-100-3	Sequence 3, Appli
7	43	41.3	974	15	US-10-101-464A-921	Sequence 921, App
8	42	40.4	184	15	US-10-156-761-14999	Sequence 14999, A
9	42	40.4	232	15	US-10-101-464A-515	Sequence 515, App
10	42	40.4	630	15	US-10-195-117-2	Sequence 2, Appli
11	42	40.4	960	8	US-08-910-386A-18	Sequence 18, Appl
12	42	40.4	1021	15	US-10-101-464A-954	Sequence 954, App
13	41.5	39.9	446	10	US-09-738-626-6262	Sequence 6262, Ap
14	41	39.4	261	15	US-10-156-761-9294	Sequence 9294, Ap
15	41	39.4	653	11	US-09-820-843A-26	Sequence 26, Appl
16	41	39.4	4123	15	US-10-213-509-5	Sequence 5, Appli
17	40	38.5	28	9	US-09-864-761-34511	Sequence 34511, A
18	40	38.5	94	15	US-10-102-806-581	Sequence 581, App
19	40	38.5	124	16	US-10-080-170-376	Sequence 376, App
20	40	38.5	142	9	US-09-908-711-105	Sequence 105, App
21	40	38.5	142	11	US-09-764-891-3245	Sequence 3245, Ap
22	40	38.5	173	15	US-10-156-761-12492	Sequence 12492, A
23	40	38.5	199	10	US-09-921-640-6	Sequence 6, Appli
24	40	38.5	258	10	US-09-921-640-9	Sequence 9, Appli
25	40	38.5	294	15	US-10-102-806-595	Sequence 595, App
26	40	38.5	394	15	US-10-156-761-8790	Sequence 8790, Ap
27	40	38.5	437	15	US-10-156-761-10200	Sequence 10200, A
28	40	38.5	907	14	US-10-008-739A-2	Sequence 2, Appli
29	40	38.5	1014	9	US-09-912-020-266	Sequence 266, App
30	40	38.5	1150	10	US-09-870-122-3	Sequence 3, Appli
31	40	38.5	1161	8	US-08-910-386A-20	Sequence 20, Appl
32	40	38.5	1164	10	US-09-870-122-1	Sequence 1, Appli
33	40	38.5	1167	10	US-09-870-122-2	Sequence 2, Appli
34	40	38.5	1181	10	US-09-870-122-23	Sequence 23, Appl
35	39.5	38.0	195	15	US-10-156-761-7790	Sequence 7790, Ap
36	39	37.5	19	11	US-09-931-325A-145	Sequence 145, App
37	39	37.5	19	12	US-09-930-915A-186	Sequence 186, App
38	39	37.5	26	11	US-09-931-325A-146	Sequence 146, App
39	39	37.5	26	12	US-09-930-915A-187	Sequence 187, App
40	39	37.5	31	12	US-09-930-915A-7	Sequence 7, Appli
41	39	37.5	31	12	US-10-082-014-13	Sequence 13, Appl
42	39	37.5	118	9	US-09-864-761-33856	Sequence 33856, A
43	39	37.5	199	15	US-10-152-300-2	Sequence 2, Appli
44	39	37.5	211	10	US-09-738-626-5094	Sequence 5094, Ap
45	39	37.5	216	9	US-09-908-180-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-10-156-761-14172
; Sequence 14172, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14172
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14172

Query Match 42.3%; Score 44; DB 15; Length 240;
Best Local Similarity 75.0%; Pred. No. 45;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          4 GVRGDFGSLAPR 15
              |  ||| |||||
Db        212 GDTGDFASLAPR 223
```

RESULT 2
US-09-983-531A-4

```
; Sequence 4, Application US/09983531A
; Patent No. US20020147132A1
; GENERAL INFORMATION:
;  APPLICANT: Fujisawa, Atsuko
;  APPLICANT:  Yamakawa, Toru
;  APPLICANT:  Shirakawa, Kamon
;  APPLICANT:  Chitose, Orii
;  APPLICANT:  Ogawa, Naoki
;  TITLE OF INVENTION: Meltrins
;  FILE REFERENCE: 11-22-99 sequence submission
;  CURRENT APPLICATION NUMBER: US/09/983,531A
;  CURRENT FILING DATE:  2001-10-24
;  PRIOR APPLICATION NUMBER: JP 8-61756
;  PRIOR FILING DATE: 1996-02-23
;  PRIOR APPLICATION NUMBER: PCT/JP96/03017
;  PRIOR FILING DATE: 1996-10-17
;  NUMBER OF SEQ ID NOS: 44
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 920
;   TYPE: PRT
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: Clone:
;   OTHER INFORMATION: JM109(pBSMel-beta), mouse beta meltrin
;   NAME/KEY: Unsure
;   LOCATION: 574
;   OTHER INFORMATION: Amino acid "Xaa" is unknown
US-09-983-531A-4

  Query Match             42.3%;  Score 44;  DB 10;  Length 920;
  Best Local Similarity   38.9%;  Pred. No. 1.8e+02;
  Matches    7;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDFGSLAPRVAR 18
              || |: | :   :|| |:
Db        555 CGKGLNGQYRKCSPRDAK 572


RESULT 3
US-10-101-464A-931
; Sequence 931, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
;  APPLICANT: Strabala, Timothy
;  APPLICANT:  Nieuwenhuizen, Nicolaas
;  APPLICANT:  Higgins, Colleen M.
;  TITLE OF INVENTION: Compositions Isolated from Plant Cells
;  TITLE OF INVENTION:  and Their Use in the Modification of Plant Cell Signaling
;  FILE REFERENCE: 11000.1020c2
;  CURRENT APPLICATION NUMBER: US/10/101,464A
;  CURRENT FILING DATE:  2002-03-18
;  PRIOR APPLICATION NUMBER: 09/704,302
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 09/228,986
;  PRIOR FILING DATE: 1999-01-12
;  PRIOR APPLICATION NUMBER: 60/162,866
;  PRIOR FILING DATE: 1999-11-01
;  PRIOR APPLICATION NUMBER: PCT/US00/00724
;  PRIOR FILING DATE: 2000-01-11
;  NUMBER OF SEQ ID NOS: 989
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 931
;   LENGTH: 998
;   TYPE: PRT
;   ORGANISM: Pinus radiata
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(998)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-10-101-464A-931

  Query Match             42.3%;  Score 44;  DB 15;  Length 998;
  Best Local Similarity   38.5%;  Pred. No. 2e+02;
  Matches   10;  Conservative    3;  Mismatches    7;  Indels    6;  Gaps    1;

Qy          1 CGSG------VRGDFGSLAPRVARQL 20
              ||||      : | :| :||  |  |
Db        849 CGSGADSMTAIAGSYGYIAPEYAYTL 874


RESULT 4
US-10-207-655-186
; Sequence 186, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
;  APPLICANT: Ledbetter, Jeffrey A.
;  APPLICANT:  Hayden-Ledbetter, Martha S.
;  TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
;  FILE REFERENCE: 390069.401C1
;  CURRENT APPLICATION NUMBER: US/10/207,655
;  CURRENT FILING DATE:  2002-07-25
;  NUMBER OF SEQ ID NOS: 426
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 186
;   LENGTH: 312
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-207-655-186

  Query Match             41.3%;  Score 43;  DB 15;  Length 312;
  Best Local Similarity   61.5%;  Pred. No. 84;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDFGSLA 13
              |||| ||  | :|
Db        164 CGSGARGGDGEVA 176


RESULT 5
US-09-291-299A-3
; Sequence 3, Application US/09291299A
; Patent No. US20020076405A1
; GENERAL INFORMATION:
;  APPLICANT: Yong, Xie
;  TITLE OF INVENTION: Antigenized Antibody Vaccine for Foot-and-Mouth Disease
;  FILE REFERENCE: 09/291,299
;  CURRENT APPLICATION NUMBER: US/09/291,299A
;  CURRENT FILING DATE:  1999-04-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 476
;   TYPE: PRT
;   ORGANISM: Sus sp.
US-09-291-299A-3

  Query Match             41.3%;  Score 43;  DB 9;  Length 476;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+02;
  Matches   11;  Conservative    2;  Mismatches    7;  Indels    2;  Gaps    1;

Qy          1 CGS--GVRGDFGSLAPRVARQL 20
              ||    :|||   || :||| |
Db        115 CGKVPNLRGDLQVLAQKVARTL 136


RESULT 6
US-10-222-100-3
; Sequence 3, Application US/10222100
; Publication No. US20030059439A1
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  APPLICANT:  Revets, Hilde
;  APPLICANT:  Cornelis, Pierre
;  APPLICANT:  De Baetselier, Patrick
;  TITLE OF INVENTION: Th1 INDUCING NATURAL ADJUVANT FOR HETEROLOGOUS ANTIGENS
```

```
;  FILE REFERENCE: 2676-5449US
;  CURRENT APPLICATION NUMBER: US/10/222,100
;  CURRENT FILING DATE:  2002-08-16
;  PRIOR APPLICATION NUMBER: PCT/EP01/01673
;  PRIOR FILING DATE: 2000-02-18
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 627
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: OprI-3D-FMDV15 fusion
US-10-222-100-3

  Query Match             41.3%;  Score 43;  DB 15;  Length 627;
  Best Local Similarity   55.6%;  Pred. No. 1.8e+02;
  Matches   10;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 SGVRGDFGSLAPRVARQL 20
              | :|||   || :||| |
Db        573 SNLRGDLQVLAQKVARTL 590


RESULT 7
US-10-101-464A-921
; Sequence 921, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
;  APPLICANT: Strabala, Timothy
;  APPLICANT:  Nieuwenhuizen, Nicolaas
;  APPLICANT:  Higgins, Colleen M.
;  TITLE OF INVENTION: Compositions Isolated from Plant Cells
;  TITLE OF INVENTION:  and Their Use in the Modification of Plant Cell Signaling
;  FILE REFERENCE: 11000.1020c2
;  CURRENT APPLICATION NUMBER: US/10/101,464A
;  CURRENT FILING DATE:  2002-03-18
;  PRIOR APPLICATION NUMBER: 09/704,302
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 09/228,986
;  PRIOR FILING DATE: 1999-01-12
;  PRIOR APPLICATION NUMBER: 60/162,866
;  PRIOR FILING DATE: 1999-11-01
;  PRIOR APPLICATION NUMBER: PCT/US00/00724
;  PRIOR FILING DATE: 2000-01-11
;  NUMBER OF SEQ ID NOS: 989
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 921
;   LENGTH: 974
;   TYPE: PRT
;   ORGANISM: Eucalyptus grandis
US-10-101-464A-921

  Query Match             41.3%;  Score 43;  DB 15;  Length 974;
  Best Local Similarity   45.0%;  Pred. No. 2.8e+02;
  Matches    9;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDFGSLAPRVARQL 20
              | | | | :| :||  | |
Db        835 CMSSVAGSYGYIAPEYAYTL 854


RESULT 8
US-10-156-761-14999
; Sequence 14999, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14999
;   LENGTH: 184
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-14999

  Query Match             40.4%;  Score 42;  DB 15;  Length 184;
  Best Local Similarity   50.0%;  Pred. No. 70;
  Matches    7;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          7 GDFGSLAPRVARQL 20
              |: |:::|||| |:
Db        108 GELGAVSPRVAAQV 121


RESULT 9
US-10-101-464A-515
; Sequence 515, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
;  APPLICANT: Strabala, Timothy
;  APPLICANT:  Nieuwenhuizen, Nicolaas
;  APPLICANT:  Higgins, Colleen M.
;  TITLE OF INVENTION: Compositions Isolated from Plant Cells
;  TITLE OF INVENTION:  and Their Use in the Modification of Plant Cell Signaling
;  FILE REFERENCE: 11000.1020c2
;  CURRENT APPLICATION NUMBER: US/10/101,464A
;  CURRENT FILING DATE:  2002-03-18
;  PRIOR APPLICATION NUMBER: 09/704,302
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 09/228,986
;  PRIOR FILING DATE: 1999-01-12
;  PRIOR APPLICATION NUMBER: 60/162,866
;  PRIOR FILING DATE: 1999-11-01
;  PRIOR APPLICATION NUMBER: PCT/US00/00724
;  PRIOR FILING DATE: 2000-01-11
;  NUMBER OF SEQ ID NOS: 989
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 515
;   LENGTH: 232
;   TYPE: PRT
;   ORGANISM: Eucalyptus grandis
US-10-101-464A-515

  Query Match             40.4%;  Score 42;  DB 15;  Length 232;
  Best Local Similarity   40.0%;  Pred. No. 89;
  Matches    8;  Conservative    3;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDFGSLAPRVARQL 20
              | | : | :| :||  |  |
Db        204 CMSAIAGSYGYIAPEYAYTL 223


RESULT 10
US-10-195-117-2
; Sequence 2, Application US/10195117
; Publication No. US20030092083A1
; GENERAL INFORMATION:
;  APPLICANT: In2Gen Co., Ltd.
;  APPLICANT:  Jeoung, Doo-il
;  APPLICANT:  Cho, Bomsoo
;  APPLICANT:  Lim, Yoon
;  APPLICANT:  Park, Saeyoung
```

```
;  APPLICANT:  Lee, Daeyeon
;  APPLICANT:  Bang, Yung-Jue
;  APPLICANT:  Yang, Hankwang
;  APPLICANT:  Kim, Dae-Kee
;  TITLE OF INVENTION: CAGE Antigen
;  FILE REFERENCE: 59258-00002
;  CURRENT APPLICATION NUMBER: US/10/195,117
;  CURRENT FILING DATE:  2002-07-11
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 630
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-195-117-2

  Query Match             40.4%;  Score 42;  DB 15;  Length 630;
  Best Local Similarity   56.2%;  Pred. No. 2.5e+02;
  Matches    9;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVA 17
              |||  | ||   || |
Db         29 GSGWSGPFGHQGPRAA 44


RESULT 11
US-08-910-386A-18
; Sequence 18, Application US/08910386A
; Publication No. US20020092041A1
;  GENERAL INFORMATION:
;    APPLICANT:  Ronald, Pamela C.
;    APPLICANT:  Wang, Guo-Liang
;    APPLICANT:  Song, Wen-Yuang
;    APPLICANT:  Hulbert, Scot
;    APPLICANT:  Richter, Todd
;    TITLE OF INVENTION:  Procedures and Materials for Conferring
;    TITLE OF INVENTION:  Disease Resisitance in Plants
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/910,386A
;      FILING DATE:  13-AUG-1997
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bastian, Kevin L.
;      REGISTRATION NUMBER:  34,774
;      REFERENCE/DOCKET NUMBER:  023070-058950US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  960 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-910-386A-18

  Query Match             40.4%;  Score 42;  DB 8;  Length 960;
  Best Local Similarity   40.0%;  Pred. No. 3.9e+02;
  Matches    8;  Conservative    3;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDFGSLAPRVARQL 20
              | | : | :| :||  |  |
Db        835 CMSAIAGSYGYIAPEYAYTL 854


RESULT 12
US-10-101-464A-954
; Sequence 954, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
;  APPLICANT: Strabala, Timothy
;  APPLICANT:  Nieuwenhuizen, Nicolaas
;  APPLICANT:  Higgins, Colleen M.
;  TITLE OF INVENTION: Compositions Isolated from Plant Cells
;  TITLE OF INVENTION:  and Their Use in the Modification of Plant Cell Signaling
;  FILE REFERENCE: 11000.1020c2
;  CURRENT APPLICATION NUMBER: US/10/101,464A
;  CURRENT FILING DATE:  2002-03-18
;  PRIOR APPLICATION NUMBER: 09/704,302
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 09/228,986
;  PRIOR FILING DATE: 1999-01-12
;  PRIOR APPLICATION NUMBER: 60/162,866
;  PRIOR FILING DATE: 1999-11-01
;  PRIOR APPLICATION NUMBER: PCT/US00/00724
;  PRIOR FILING DATE: 2000-01-11
;  NUMBER OF SEQ ID NOS: 989
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 954
;   LENGTH: 1021
;   TYPE: PRT
;   ORGANISM: Eucalyptus grandis
US-10-101-464A-954

  Query Match             40.4%;  Score 42;  DB 15;  Length 1021;
  Best Local Similarity   40.0%;  Pred. No. 4.2e+02;
  Matches    8;  Conservative    3;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDFGSLAPRVARQL 20
              | | : | :| :||  |  |
Db        861 CMSAIAGSYGYIAPEYAYTL 880


RESULT 13
US-09-738-626-6262
; Sequence 6262, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6262
;   LENGTH: 446
;   TYPE: PRT
```

```
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-6262

  Query Match            39.9%;  Score 41.5;  DB 10;  Length 446;
  Best Local Similarity   55.0%;  Pred. No. 2.1e+02;
  Matches   11;  Conservative    2;  Mismatches    4;  Indels    3;  Gaps    1;

Qy          4 GVR---GDFGSLAPRVARQL 20
              |||   || | || :| :||
Db        261 GVRIDSGDLGVLARKVRKQL 280


RESULT 14
US-10-156-761-9294
; Sequence 9294, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9294
;   LENGTH: 261
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-9294

  Query Match            39.4%;  Score 41;  DB 15;  Length 261;
  Best Local Similarity   52.6%;  Pred. No. 1.4e+02;
  Matches   10;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLAPRVARQL 20
              |  |  | |: | |:||||
Db         97 GQSVLLDIGTTAYRLARQL 115


RESULT 15
US-09-820-843A-26
; Sequence 26, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
;  APPLICANT: Council of Scientific and Industrial Research
;  TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEI
;  TITLE OF INVENTION:  USEFUL AS ANTI-INFECTIVES
;  FILE REFERENCE: Q63915
;  CURRENT APPLICATION NUMBER: US/09/820,843A
;  CURRENT FILING DATE:  2001-03-30
;  NUMBER OF SEQ ID NOS: 118
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
;   LENGTH: 653
;   TYPE: PRT
;   ORGANISM: Vibrio cholerae
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: iron(III) ABC transporter, permease protein
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: gi|9654609
US-09-820-843A-26

  Query Match            39.4%;  Score 41;  DB 11;  Length 653;
  Best Local Similarity   52.9%;  Pred. No. 3.8e+02;
  Matches    9;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 GVRGDFGSLAPRVARQL 20
              || |  | | | :|| |
Db        249 GVIGFIGLLTPNIARSL 265


Search completed: October  9, 2003, 10:20:48
Job time : 22.3333 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:02:18 ; Search time 58.3333 Seconds
(without alignments)
54.421 Million cell updates/sec

Title: US-09-549-186B-8
Perfect score: 101
Sequence: 1 CGSGVRGDFGSLALRVARQL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	92	91.1	19	24	ABU08643	Foot and mouth dis
2	85	84.2	19	24	ABU08642	Foot and mouth dis
3	85	84.2	20	21	AAY94587	Envelope gene epit
4	85	84.2	31	4	AAP30107	Sequence of VP1 ca
5	85	84.2	31	6	AAP50728	Foot-and-mouth dis
6	85	84.2	31	18	AAW17082	Foot and mouth dis
7	85	84.2	35	21	AAY68501	Target antigenic p
8	85	84.2	75	21	AAY68528	Synthetic foot and
9	85	84.2	216	4	AAP30205	Sequence encoded b
10	85	84.2	220	4	AAP30195	Sequence encoded b
11	85	84.2	233	4	AAP30204	Sequence encoded b
12	85	84.2	405	4	AAP30202	Sequence encoded b
13	85	84.2	406	4	AAP30203	Sequence encoded b
14	85	84.2	609	4	AAP30206	Sequence encoded b
15	84	83.2	20	14	AAR38543	Foot and Mouth Dis
16	84	83.2	20	14	AAR41793	Foot and Mouth Dis
17	83	82.2	31	18	AAW17085	Foot and mouth dis
18	82	81.2	31	18	AAW17086	Foot and mouth dis
19	81	80.2	31	18	AAW17087	Foot and mouth dis
20	79	78.2	31	18	AAW17084	Foot and mouth dis
21	78	77.2	31	18	AAW17083	Foot and mouth dis
22	78	77.2	31	18	AAW17088	Foot and mouth dis
23	77	76.2	35	21	AAY68502	Target antigenic p
24	77	76.2	46	21	AAY68503	Target antigenic p
25	77	76.2	65	21	AAY68527	Synthetic foot and
26	77	76.2	75	21	AAY68529	Synthetic foot and
27	75	74.3	35	7	AAP60895	Peptide containing
28	75	74.3	35	9	AAP82746	VP1(A12,119) contg
29	71	70.3	35	21	AAY68508	Consensus antigeni
30	71	70.3	213	4	AAP30198	Sequence encoded b
31	71	70.3	233	4	AAP30209	Sequence of antige
32	70	69.3	18	4	AAP30110	Sequence of VP1 ca
33	67	66.3	20	5	AAP40103	Sequence at antige
34	67	66.3	216	5	AAP40085	Sequence of foot a
35	64	63.4	248	4	AAP30207	Sequence of an FMD
36	63	62.4	35	21	AAY68509	Consensus antigeni
37	63	62.4	55	21	AAY68533	Synthetic foot and
38	58	57.4	218	4	AAP30196	Sequence encoded b
39	56	55.4	28	4	AAP30316	Sequence of VP1 ca
40	56	55.4	213	21	AAB15428	FMDV antigen for r
41	56	55.4	213	21	AAB15397	Foot-and-mouth dis
42	56	55.4	213	21	AAB15398	Foot-and-mouth dis
43	56	55.4	213	21	AAB15402	Foot-and-mouth dis
44	54	53.5	36	7	AAP60896	Peptide containing
45	54	53.5	36	9	AAP82747	VP1(A24,C) contg.

ALIGNMENTS

RESULT 1
ABU08643
ID ABU08643 standard; peptide; 19 AA.
XX
AC ABU08643;
XX
DT 04-JUN-2003 (first entry)
XX
DE Foot and mouth disease virus (FMDV) VP1-A12141-159, variant FL.
XX
KW Foot and mouth disease virus; FMDV; VP1-A12 region; immunosuppressive;
KW antiinflammatory; antiarthritic; antirheumatic; dermatological;
KW antidiabetic; antianaemic; neuroprotective; vaccine;
KW anti-metallothionein antibody; humoral immune response;
KW autoimmune disease; Grave's disease; rheumatoid arthritis;
KW systemic lupus erythematosus; Type 1 diabetes; pernicious anaemia;
KW multiple sclerosis; Sjorgen's disease.
XX
OS Foot and mouth disease virus.
XX
PN US2003007973-A1.
XX
PD 09-JAN-2003.
XX
PF 24-JUN-2002; 2002US-0178909.
XX
PR 22-JUN-2001; 2001US-300346P.
XX
PA (LYNE/) LYNES M A.
XX
PI Lynes MA;
XX

DR WPI; 2003-353082/33.
XX
PT Treating autoimmune diseases, e.g. Grave's disease, arthritis, lupus,
PT or diabetes by administering a composition consisting of an
PT anti-metallothionein antibody to stimulate a humoral immune response in
PT a subject -
XX
PS Example 7; Fig 9; 24pp; English.
XX
CC The invention describes a method of treating a subject comprising
CC administering to the subject a composition consisting of an
CC anti-metallothionein antibody to stimulate a humoral immune response in
CC a subject. The methods and compositions are useful for treating
CC autoimmune diseases, e.g. Grave's disease, rheumatoid arthritis,
CC systemic lupus erythematosus, Type 1 diabetes, pernicious anaemia,
CC multiple sclerosis or Sjorgen's disease. This is the amino acid
CC sequence of foot and mouth disease virus (FMDV) VP1-A12141-159 region
CC used as an immunogen.
XX
SQ Sequence 19 AA;

Query Match 91.1%; Score 92; DB 24; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.6e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GSGVRGDFGSLALRVARQL 20
||||||||||||||||||||
Db 1 GSGVRGDFGSLALRVARQL 19

RESULT 2
ABU08642
ID ABU08642 standard; peptide; 19 AA.
XX
AC ABU08642;
XX
DT 04-JUN-2003 (first entry)
XX
DE Foot and mouth disease virus (FMDV) VP1-A12141-159, variant FP.
XX
KW Foot and mouth disease virus; FMDV; VP1-A12 region; immunosuppressive;
KW antiinflammatory; antiarthritic; antirheumatic; dermatological;
KW antidiabetic; antianaemic; neuroprotective; vaccine;
KW anti-metallothionein antibody; humoral immune response;
KW autoimmune disease; Grave's disease; rheumatoid arthritis;
KW systemic lupus erythematosus; Type 1 diabetes; pernicious anaemia;
KW multiple sclerosis; Sjorgen's disease.
XX
OS Foot and mouth disease virus.
XX
PN US2003007973-A1.
XX
PD 09-JAN-2003.
XX
PF 24-JUN-2002; 2002US-0178909.
XX
PR 22-JUN-2001; 2001US-300346P.
XX
PA (LYNE/) LYNES M A.
XX
PI Lynes MA;
XX
DR WPI; 2003-353082/33.
XX
PT Treating autoimmune diseases, e.g. Grave's disease, arthritis, lupus,
PT or diabetes by administering a composition consisting of an
PT anti-metallothionein antibody to stimulate a humoral immune response in
PT a subject -
XX
PS Example 7; Fig 9; 24pp; English.
XX
CC The invention describes a method of treating a subject comprising
CC administering to the subject a composition consisting of an
CC anti-metallothionein antibody to stimulate a humoral immune response in
CC a subject. The methods and compositions are useful for treating
CC autoimmune diseases, e.g. Grave's disease, rheumatoid arthritis,
CC systemic lupus erythematosus, Type 1 diabetes, pernicious anaemia,
CC multiple sclerosis or Sjorgen's disease. This is the amino acid
CC sequence of foot and mouth disease virus (FMDV) VP1-A12141-159 region
CC used as an immunogen.
XX
SQ Sequence 19 AA;

Query Match 84.2%; Score 85; DB 24; Length 19;
Best Local Similarity 94.7%; Pred. No. 4.2e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GSGVRGDFGSLALRVARQL 20
|||||||||||| ||||||
Db 1 GSGVRGDFGSLAPRVARQL 19

RESULT 3
AAY94587
ID AAY94587 standard; Peptide; 20 AA.
XX
AC AAY94587;
XX
DT 10-JAN-2001 (first entry)
XX
DE Envelope gene epitope of foot and mouth disease virus.
XX
KW Hepatitis B virus nucleocapsid antigen; HBcAg; T cell epitope;
KW cytolytic T lymphocyte; immunogenic; ICE; CTL; FMDV;
KW immunodominant core epitope; immunisation; envelope gene.
XX
OS Foot and mouth disease virus.
XX
PN WO200026385-A1.
XX
PD 11-MAY-2000.
XX
PF 05-NOV-1999; 99WO-US26291.
XX
PR 05-NOV-1998; 98US-0107169.
XX
PA (POWD-) POWDERJECT VACCINES INC.
XX
PI Fuller DL, Fuller JT;
XX
DR WPI; 2000-451623/39.
XX
PT Use of expression vector for nucleic acid immunization that comprises
PT promoter and recombinant nucleic acid sequences encoding Hepatitis B
PT core antigen and T cell epitope from antigen -
XX
PS Example 6; Page 39; 55pp; English.
XX
CC The present invention relates to an immunogeneic recombinant
CC nucleic acid molecule. The molecule consists of a modified hepatitis
CC B virus nucleocapsid antigen (HBcAg) with a T cell epitope sequence
CC inserted within the HBcAg. The creation of a unique restriction site
CC in HBcAg facilitated the insertion of the T cell epitope into the DNA
CC encoding the immunodominant core epitope of the HBcAg. An example of a
CC suitable insertion epitope is the present sequence, the
CC neutralisation epitope from the foot and mouth disease virus envelope
CC gene. Alternatively other T cell epitopes may be inserted
CC (AAY94583, AAY94584, AAY94585, AAY94586, AAY94588). The recombinant
CC nucleic acid molecule may then be used as a reagent in various nucleic
CC acid immunisation strategies. The advantage of this method of
CC immunisation is that the nucleic acid reagents that encode hybrid HBcAg
CC generate an extremely high frequency cellular immune response against
CC the CTL epitope.
XX
SQ Sequence 20 AA;

Query Match 84.2%; Score 85; DB 21; Length 20;
Best Local Similarity 94.7%; Pred. No. 4.5e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 GSGVRGDFGSLALRVARQL 20
              |||||||||||| ||||||
Db          1 GSGVRGDFGSLAPRVARQL 19
```

RESULT 4
AAP30107
ID AAP30107 standard; Peptide; 31 AA.
XX
AC AAP30107;
XX
DT 03-APR-1992 (first entry)
XX
DE Sequence of VP1 capisid protein residues 130-160 from the amino-
DE terminus, FMDV, Tubingen type A, subtype 12, strain 119.
XX
KW Antigen; Picornavirus; capsid protein; antibody; detection;
KW vaccine; diagnosis.
XX
OS Foot and mouth disease virus.
XX
PN WO8303547-A.
XX
PD 27-OCT-1983.
XX
PF 14-APR-1983; 83WO-0002644.
XX
PR 25-MAR-1983; 83US-0478847.
PR 14-APR-1982; 82US-0368308.
PR 20-SEP-1984; 84US-0653475.
PR 18-DEC-1984; 84US-0682819.
XX
PA (BITT/) BITTLE J L.
PA (SCRI-) SCRIPPS CLINIC & RE.
XX
PI Bittle JL, Lerner RA;
XX
DR WPI; 1983-807942/44.
XX
PT Antigenic peptide(s) corresp. to picornavirus capsid protein -
PT useful in prodn. of vaccines and in diagnostic tests
XX
PS Example; Page 26; 90pp; English.
XX
CC The peptides of the invention corresp. to a region on the antigenic
CC Picornavirus capsid protein. The capsid protein FMDV VP1 or polio
CC virus VP1. When linked to carriers the peptides are immunogenic.
CC Dose is 20 ug-2mg peptide for inoculations.
XX
SQ Sequence 31 AA;

Query Match 84.2%; Score 85; DB 4; Length 31;
Best Local Similarity 94.7%; Pred. No. 6.8e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 GSGVRGDFGSLALRVARQL 20
              |||||||||||| ||||||
Db         11 GSGVRGDFGSLAPRVARQL 29
```

RESULT 5
AAP50728
ID AAP50728 standard; protein; 31 AA.
XX
AC AAP50728;
XX
DT 23-OCT-1991 (first entry)
XX
DE Foot-and-mouth disease antigenic peptide.
XX
KW Foot-and-mouth disease; vaccine; antigen;
XX
OS Foot-and-mouth disease virus.
XX
PN US4544500-A.
XX
PD 01-OCT-1985.
XX
PF 18-DEC-1984; 84US-0682819.
XX
PR 18-DEC-1984; 84US-0682819.
XX
PA (SCRI-) SCRIPPS CLINIC & RESEARCH FOUNDATION.
XX
PI Bittle JL, Lerner RA;
XX
DR WPI; 1985-262823/42.
XX
PT Synthetic foot and mouth disease antigen - comprising 20 amino acid
PT peptide corresp. to virus protein VP1.
XX
PS Disclosure; page 4; 7pp; English.
XX
CC The peptide corresponds to amino acids 130-160 of the foot and mouth
CC disease virus VP1 protein. It represents a monospecific synthetic
CC antigenic determinant. The peptide can be used to vaccinate animals,
CC esp. swine and cattle, against the virus with a single inoculation.
XX
SQ Sequence 31 AA;

Query Match 84.2%; Score 85; DB 6; Length 31;
Best Local Similarity 94.7%; Pred. No. 6.8e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 GSGVRGDFGSLALRVARQL 20
              |||||||||||| ||||||
Db         11 GSGVRGDFGSLAPRVARQL 29
```

RESULT 6
AAW17082
ID AAW17082 standard; peptide; 31 AA.
XX
AC AAW17082;
XX
DT 13-JUN-1997 (first entry)
XX
DE Foot and mouth disease virus G-H loop of protein VP-1.
XX
KW Foot and mouth disease; aphthovirus; vaccine; antigen; epitope;
KW immunogen; G-H loop; VP-1; deletion; mutant; binding site removal;
KW antibody; non-infectious; attenuated.
XX
OS Foot and mouth disease virus.
XX
PN US5612040-A.
XX
PD 18-MAR-1997.
XX
PF 07-APR-1995; 95US-0418716.
XX
PR 07-APR-1995; 95US-0418716.
XX
PA (USDA) US SEC OF AGRIC.
XX
PI Baxt B, Berinstein A, Kang AS, Mason PW, Reider E;
XX
DR WPI; 1997-192081/17.
XX
PT Genetically modified foot-and-mouth disease virus - lacks
PT cell-binding site and is non-infectious, useful in vaccines
XX

```
PS   Disclosure; Figure 1; 10pp; English.
XX
CC   AAW17082 represents amino acids 130-163 of wild-type foot and mouth
CC   disease (FMD) virus  VP-1 structural protein. The sequence represents
CC   the G-H loop, a flexible loop between the G and H beta strands of VP-1.
CC   The G-H loop is the main antigenic site of VP-1 and is also involved in
CC   cell binding. The FMD virus G-H loop peptide can be modified in a FMD
CC   virus by deletion of the amino acid (aa) sequence
CC   Gly-Val-Arg-Gly-Asp-Phe (aa 142-147) and replacing this sequence with the
CC   sequence Asn-Pro. This results in the virus retaining its immunogenicity
CC   but is not infectious because the cell binding site has been removed. The
CC   genetically modified FMD virus is useful in vaccines against the
CC   disease.
XX
SQ   Sequence   31 AA;

  Query Match             84.2%;  Score 85;  DB 18;  Length 31;
  Best Local Similarity   94.7%;  Pred. No. 6.8e-06;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              ||||||||||||| ||||||
Db         11 GSGVRGDFGSLAPRVARQL 29


RESULT 7
AAY68501
ID   AAY68501 standard; peptide; 35 AA.
XX
AC   AAY68501;
XX
DT   05-MAY-2000  (first entry)
XX
DE   Target antigenic peptide derived from amino acids 134-169 of VP1.
XX
KW   Antigenic peptide; VP1 capsid protein; FMDV strain A12; epitope;
KW   helper T-cell epitope; immune response; vaccine; FMDV infection.
XX
OS   Foot and mouth disease virus.
XX
PN   WO9966954-A1.
XX
PD   29-DEC-1999.
XX
PF   21-JUN-1999;   99WO-US13921.
XX
PR   20-JUN-1998;   98US-0100600.
XX
PA   (UNBI-) UNITED BIOMEDICAL INC.
XX
PI   Wang CY,  Shen M;
XX
DR   WPI; 2000-160563/14.
XX
PT   Synthetic peptide used in protecting animals against Foot-and-Mouth
PT   Disease Virus infections comprises a  helper T-cell epitope and a
PT   Foot-and-Mouth Disease Virus epitope conjugate -
XX
PS   Claim 1; Page 45; 115pp; English.
XX
CC   The present sequence represents a target antigenic peptide derived from
CC   the VP1 capsid protein of Foot and Mouth Disease Virus (FMDV) strain A12.
CC   The peptide is used to design synthetic epitopes, which are conjugated
CC   to a helper T-cell epitope and optionally an immunostimulator sequence.
CC   The peptides are covalently linked to each other, preferably by peptide
CC   bonds generated through direct synthesis. The synthetic vaccines of the
CC   invention require lower levels of biosecurity, have a reduced loss of
CC   immunogenic activity, are less prone to product instability and
CC   lot-to-lot variability and do not need periodic revision due to
CC   antigenic variation in the field. The peptides are used as immunogens
CC   in vaccines to generate an enhanced immune response against a FMDV
CC   antigen in animals, such as swine, cattle, sheep, goats and susceptible
CC   wild species. They can also be used to diagnose FMDV infection in a
CC   mammal.
XX
SQ   Sequence   35 AA;

  Query Match             84.2%;  Score 85;  DB 21;  Length 35;
  Best Local Similarity   94.7%;  Pred. No. 7.6e-06;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              ||||||||||||| ||||||
Db          7 GSGVRGDFGSLAPRVARQL 25


RESULT 8
AAY68528
ID   AAY68528 standard; peptide; 75 AA.
XX
AC   AAY68528;
XX
DT   05-MAY-2000  (first entry)
XX
DE   Synthetic foot and mouth disease virus immunogen.
XX
KW   Antigenic peptide; VP1 capsid protein; serotype Asia; epitope;
KW   helper T-cell epitope; immune response; vaccine; FMDV infection.
XX
OS   Synthetic.
OS   Yersinia sp.
OS   Foot and mouth disease virus.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..16
FT                   /note= "Yersinia invasin domain (immunostimulator
FT                    peptide"
FT   Peptide         17..18
FT                   /note= "spacer"
FT   Peptide         19..38
FT                   /note= "autologous helper T-cell epitope derived
FT                    from VP1 capsid protein"
FT   Peptide         39..40
FT                   /note= "spacer"
FT   Peptide         41..75
FT                   /note= "synthetic VP1 capsid protein epitope"
XX
PN   WO9966954-A1.
XX
PD   29-DEC-1999.
XX
PF   21-JUN-1999;   99WO-US13921.
XX
PR   20-JUN-1998;   98US-0100600.
XX
PA   (UNBI-) UNITED BIOMEDICAL INC.
XX
PI   Wang CY,  Shen M;
XX
DR   WPI; 2000-160563/14.
XX
PT   Synthetic peptide used in protecting animals against Foot-and-Mouth
PT   Disease Virus infections comprises a  helper T-cell epitope and a
PT   Foot-and-Mouth Disease Virus epitope conjugate -
XX
PS   Claim 8; Page 96; 115pp; English.
XX
CC   The present sequence represents a foot and mouth disease (FMDV) vaccine
CC   of the invention. The specification describes epitopes derived from an
CC   antigen of the VP1 capsid protein of FMDV. The VP1 epitopes are
CC   conjugated to a helper T-cell epitope and optionally an immunostimulator
CC   sequence. The peptides are covalently linked to each other, preferably
CC   by peptide bonds generated through direct synthesis. The synthetic
CC   vaccines of the invention require lower levels of biosecurity, have a
CC   reduced loss of immunogenic activity, are less prone to product
CC   instability and lot-to-lot variability and do not need periodic revision
```

```
CC   due to antigenic variation in the field. The peptides are used as
CC   immunogens in vaccines to generate an enhanced immune response against
CC   a FMDV antigen in animals, such as swine, cattle, sheep, goats and
CC   susceptible wild species. They can also be used to diagnose FMDV
CC   infection in a mammal.
XX
SQ   Sequence   75 AA;

  Query Match             84.2%;  Score 85;  DB 21;  Length 75;
  Best Local Similarity   94.7%;  Pred. No. 1.6e-05;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              |||||||||||| ||||||
Db         47 GSGVRGDFGSLAPRVARQL 65


RESULT 9
AAP30205
ID   AAP30205 standard; Protein; 216 AA.
XX
AC   AAP30205;
XX
DT   29-JUL-1992  (first entry)
XX
DE   Sequence encoded by direct expression vector pFM 10 which
DE   comprises the promoter and operator of the E. coli tryptophan
DE   operon AAs 1-211 of the VP3 gene linked to 4 AAs from pBR322.
XX
KW   Vaccine; immunogen; antigen; viral protein.
XX
OS   Foot and mouth disease virus.
XX
PN   EP68693-A.
XX
PD   05-JAN-1983.
XX
PF   11-JUN-1982;   82EP-0303040.
XX
PR   04-MAY-1982;   82US-0374855.
PR   16-JUN-1981;   81US-0274103.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Kleid DG,  Yansura DG;
XX
DR   WPI; 1983-05055K/03.
DR   N-PSDB; AAN30144.
XX
PT   Polypeptide antigens of foot-and-mouth disease - obtd. by
PT   recombinant DNA technology for vaccine prodn.
XX
PS   Example; Page 36-37; 81pp; English.
XX
CC   The inventors claim a polypeptide comprising at least one antigen of
CC   FMDV, or a fusion polypeptide of at least two antigens from at least
CC   two different FMDV strands, and encoding DNA.  Pref. the fusion
CC   polypeptide has the sequence of two of the A, C, O, Asia-1, SAT 1,
CC   SAT 2 and SAT 3 type strains of the virus.  The polypeptides pref.
CC   comprise AAs 55-177 or 1-210 of the VP3 virus.  Fusion polypeptides
CC   may comprise AAs 1-210 fused to 8-210 of VP3 virus protein.  pFM1,
CC   pFM2, pFM3, pFM10, pFM20, pFMB1, pFMC, pFMD, pFMF and pFMG are also
CC   claimed.
XX
SQ   Sequence   216 AA;

  Query Match             84.2%;  Score 85;  DB 4;  Length 216;
  Best Local Similarity   94.7%;  Pred. No. 4.4e-05;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              |||||||||||| ||||||
Db        141 GSGVRGDFGSLAPRVARQL 159


RESULT 10
AAP30195
ID   AAP30195 standard; Protein; 220 AA.
XX
AC   AAP30195;
XX
DT   29-JUL-1992  (first entry)
XX
DE   Sequence encoded by VP3 gene of FMDV type A12 (T465).
XX
KW   Vaccine; immunogen; antigen; viral protein.
XX
OS   Foot and mouth disease virus.
XX
PN   EP68693-A.
XX
PD   05-JAN-1983.
XX
PF   11-JUN-1982;   82EP-0303040.
XX
PR   04-MAY-1982;   82US-0374855.
PR   16-JUN-1981;   81US-0274103.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Kleid DG,  Yansura DG;
XX
DR   WPI; 1983-05055K/03.
DR   N-PSDB; AAN30127.
XX
PT   Polypeptide antigens of foot-and-mouth disease - obtd. by
PT   recombinant DNA technology for vaccine prodn.
XX
PS   Disclosure; Fig 2; 81pp; English.
XX
CC   The inventors claim a polypeptide comprising at least one antigen of
CC   FMDV, or a fusion polypeptide of at least two antigens from at least
CC   two different FMDV strands, and encoding DNA.  Pref. the fusion
CC   polypeptide has the sequence of two of the A, C, O, Asia-1, SAT 1,
CC   SAT 2 and SAT 3 type strains of the virus.  The polypeptides pref.
CC   comprise AAs 55-177 or 1-210 of the VP3 virus.  Fusion polypeptides
CC   may comprise AAs 1-210 fused to 8-210 of VP3 virus protein.  pFM1,
CC   pFM2, pFM3, pFM10, pFM20, pFMB1, pFMC, pFMD, pFMF and pFMG are also
CC   claimed.
XX
SQ   Sequence   220 AA;

  Query Match             84.2%;  Score 85;  DB 4;  Length 220;
  Best Local Similarity   94.7%;  Pred. No. 4.4e-05;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              |||||||||||| ||||||
Db        140 GSGVRGDFGSLAPRVARQL 158


RESULT 11
AAP30204
ID   AAP30204 standard; Protein; 233 AA.
XX
AC   AAP30204;
XX
DT   29-JUL-1992  (first entry)
XX
DE   Sequence encoded by pFM3 which comprises 17 AAs from the LE' gene
DE   construction (fragment 3a), linked to a methionine AA (CNBr cleavable),
DE   linked to the AAs 1-211 of the VP3 FMDV A12, linked to 4 AAs from
DE   pBR322.
XX
KW   Vaccine; immunogen; antigen; viral protein.
XX
```

OS Foot and mouth disease virus and Escherichia coli.
XX
PN EP63693-A.
XX
PD 05-JAN-1983.
XX
PF 11-JUN-1982; 82EP-0303040.
XX
PR 04-MAY-1982; 82US-0374855.
PR 16-JUN-1981; 81US-0274103.
XX
PA (GETH) GENENTECH INC.
XX
PI Kleid DG, Yansura DG;
XX
DR WPI; 1983-05055K/03.
DR N-PSDB; AAN30142.
XX
PT Polypeptide antigens of foot-and-mouth disease - obtd. by
PT recombinant DNA technology for vaccine prodn.
XX
PS Example; Page 33-34; 81pp; English.
XX
CC The inventors claim a polypeptide comprising at least one antigen of
CC FMDV, or a fusion polypeptide of at least two antigens from at least
CC two different FMDV strands, and encoding DNA. Pref. the fusion
CC polypeptide has the sequence of two of the A, C, O, Asia-1, SAT 1,
CC SAT 2 and SAT 3 type strains of the virus. The polypeptides pref.
CC comprise AAs 55-177 or 1-210 of the VP3 virus. Fusion polypeptides
CC may comprise AAs 1-210 fused to 8-210 of VP3 virus protein. pFM1,
CC pFM2, pFM3, pFM10, pFM20, pFMB1, pFMC, pFMD, pFMF and pFMG are also
CC claimed.
XX
SQ Sequence 233 AA;

Query Match 84.2%; Score 85; DB 4; Length 233;
Best Local Similarity 94.7%; Pred. No. 4.7e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GSGVRGDFGSLALRVARQL 20
||||||||||||| ||||||
Db 158 GSGVRGDFGSLAPRVARQL 176

RESULT 12
AAP30202
ID AAP30202 standard; Protein; 405 AA.
XX
AC AAP30202;
XX
DT 29-JUL-1992 (first entry)
XX
DE Sequence encoded by pFM1 which comprises 190 AAs form the LE' gene
DE construction linked to 6 AAs from the linker molecule (fragment 4),
DE linked to 205 AAs from the VP3 FMDV A12 (codons 7-211), linked to
DE 4 AAs from pBR322.
XX
KW Vaccine; immunogen; antigen; viral protein.
XX
OS Foot and mouth disease virus and Escherichia coli.
XX
PN EP68693-A.
XX
PD 05-JAN-1983.
XX
PF 11-JUN-1982; 82EP-0303040.
XX
PR 04-MAY-1982; 82US-0374855.
PR 16-JUN-1981; 81US-0274103.
XX
PA (GETH) GENENTECH INC.
XX
PI Kleid DG, Yansura DG;
XX
DR WPI; 1983-05055K/03.
DR N-PSDB; AAN30140.
XX
PT Polypeptide antigens of foot-and-mouth disease - obtd. by
PT recombinant DNA technology for vaccine prodn.
XX
PS Example; Page 29-31; 81pp; English.
XX
CC The inventors claim a polypeptide comprising at least one antigen of
CC FMDV, or a fusion polypeptide of at least two antigens from at least
CC two different FMDV strands, and encoding DNA. Pref. the fusion
CC polypeptide has the sequence of two of the A, C, O, Asia-1, SAT 1,
CC SAT 2 and SAT 3 type strains of the virus. The polypeptides pref.
CC comprise AAs 55-177 or 1-210 of the VP3 virus. Fusion polypeptides
CC may comprise AAs 1-210 fused to 8-210 of VP3 virus protein. pFM1,
CC pFM2, pFM3, pFM10, pFM20, pFMB1, pFMC, pFMD, pFMF and pFMG are also
CC claimed.
XX
SQ Sequence 405 AA;

Query Match 84.2%; Score 85; DB 4; Length 405;
Best Local Similarity 94.7%; Pred. No. 7.9e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GSGVRGDFGSLALRVARQL 20
||||||||||||| ||||||
Db 330 GSGVRGDFGSLAPRVARQL 348

RESULT 13
AAP30203
ID AAP30203 standard; Protein; 406 AA.
XX
AC AAP30203;
XX
DT 29-JUL-1992 (first entry)
XX
DE Sequence encoded by pFM2 which comprises190 AAs from the LE' gene,
DE linked to a methionine AA (CNBr cleavable), linked to the AAs 1-211
DE of the VP3 FMDV A12 (codons 1-7 are derived form synthetic DNA, 8-211
DE are natural codons), linked to 4 amino acids from pBR322.
XX
KW Vaccine; immunogen; antigen; viral protein.
XX
OS Foot and mouth disease virus and Escherichia coli.
XX
PN EP68693-A.
XX
PD 05-JAN-1983.
XX
PF 11-JUN-1982; 82EP-0303040.
XX
PR 04-MAY-1982; 82US-0374855.
PR 16-JUN-1981; 81US-0274103.
XX
PA (GETH) GENENTECH INC.
XX
PI Kleid DG, Yansura DG;
XX
DR WPI; 1983-05055K/03.
DR N-PSDB; AAN30141.
XX
PT Polypeptide antigens of foot-and-mouth disease - obtd. by
PT recombinant DNA technology for vaccine prodn.
XX
PS Example; Page 31-32; 81pp; English.
XX
CC The inventors claim a polypeptide comprising at least one antigen of
CC FMDV, or a fusion polypeptide of at least two antigens from at least
CC two different FMDV strands, and encoding DNA. Pref. the fusion
CC polypeptide has the sequence of two of the A, C, O, Asia-1, SAT 1,
CC SAT 2 and SAT 3 type strains of the virus. The polypeptides pref.

```
CC   comprise AAs 55-177 or 1-210 of the VP3 virus.  Fusion polypeptides
CC   may comprise AAs 1-210 fused to 8-210 of VP3 virus protein.  pFM1,
CC   pFM2, pFM3, pFM10, pFM20, pFMB1, pFMC, pFMD, pFMF and pFMG are also
CC   claimed.
XX
SQ   Sequence   406 AA;

  Query Match             84.2%;  Score 85;  DB 4;  Length 406;
  Best Local Similarity   94.7%;  Pred. No. 8e-05;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              |||||||||||| ||||||
Db        331 GSGVRGDFGSLAPRVARQL 349


RESULT 14
AAP30206
ID   AAP30206 standard; Protein; 609 AA.
XX
AC   AAP30206;
XX
DT   29-JUL-1992  (first entry)
XX
DE   Sequence encoded by pFM 20 which is the polyantigen comprising 190 AAs
DE   coded be the LE' gene construction, linked to a methionine AA (CNBr
DE   cleavable), linked to the AAs 1-211 of the VP3 FMDV A12 (codons 1-7
DE   are derived from synthetic DNA), linked to AAs 8-211 of the VP3
DE   FMDV A12, linked to 4 AAs from pBR322.
XX
KW   Vaccine; immunogen; antigen; viral protein.
XX
OS   Foot and mouth disease virus and Escherichia coli.
XX
PN   EP68693-A.
XX
PD   05-JAN-1983.
XX
PF   11-JUN-1982;   82EP-0303040.
XX
PR   04-MAY-1982;   82US-0374855.
PR   16-JUN-1981;   81US-0274103.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Kleid DG,  Yansura DG;
XX
DR   WPI; 1983-05055K/03.
DR   N-PSDB; AAN30145.
XX
PT   Polypeptide antigens of foot-and-mouth disease - obtd. by
PT   recombinant DNA technology for vaccine prodn.
XX
PS   Example; Page 38-41; 81pp; English.
XX
CC   The inventors claim a polypeptide comprising at least one antigen of
CC   FMDV, or a fusion polypeptide of at least two antigens from at least
CC   two different FMDV strands, and encoding DNA.  Pref. the fusion
CC   polypeptide has the sequence of two of the A, C, O, Asia-1, SAT 1,
CC   SAT 2 and SAT 3 type strains of the virus.  The polypeptides pref.
CC   comprise AAs 55-177 or 1-210 of the VP3 virus.  Fusion polypeptides
CC   may comprise AAs 1-210 fused to 8-210 of VP3 virus protein.  pFM1,
CC   pFM2, pFM3, pFM10, pFM20, pFMB1, pFMC, pFMD, pFMF and pFMG are also
CC   claimed.
XX
SQ   Sequence   609 AA;

  Query Match             84.2%;  Score 85;  DB 4;  Length 609;
  Best Local Similarity   94.7%;  Pred. No. 0.00012;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              |||||||||||| ||||||
Db        331 GSGVRGDFGSLAPRVARQL 349


RESULT 15
AAR38543
ID   AAR38543 standard; peptide; 20 AA.
XX
AC   AAR38543;
XX
DT   25-MAR-2003  (updated)
DT   11-JAN-1994  (first entry)
XX
DE   Foot and Mouth Disease Virus VPI antigenic peptide.
XX
KW   Multivalent vaccine; polylysine; homopolymer; dendritic core;
KW   multiple antigen peptide system; MAPS; Foot and Mouth Disease;
KW   FMDV; VPI protein; Aphthovirus; vaccine.
XX
OS   Foot and Mouth Disease Virus.
XX
PN   US5229490-A.
XX
PD   20-JUL-1993.
XX
PF   20-DEC-1990;   90US-0631185.
XX
PR   06-MAY-1987;   87US-0047204.
PR   30-JUN-1987;   87US-0068840.
PR   12-APR-1989;   89US-0336845.
PR   20-DEC-1990;   90US-0631185.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Tam JP;
XX
DR   WPI; 1993-242534/30.
XX
PT   Multiple antigenic peptide systems - have dendritic core attached
PT   covalently to antigens, used as vaccine
XX
PS   Claim 26; Column 24; 23pp; English.
XX
CC   This is a preferred FMDV antigenic peptide suitable for
CC   construction of a Multiple Antigen Peptide System (MAPS) by
CC   covalent coupling to a dendritic homopolymer core (esp.
CC   polylysine). The antigen is derived from the Foot and Mouth Disease
CC   Virus VPI protein.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence   20 AA;

  Query Match             83.2%;  Score 84;  DB 14;  Length 20;
  Best Local Similarity   94.7%;  Pred. No. 6.3e-06;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              ||||||| ||||||||||
Db          1 GSGVRGDSGSLALRVARQL 19


Search completed: October  9, 2003, 10:14:18
Job time : 58.3333 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:10:09 ; Search time 19.6667 Seconds
(without alignments)
97.799 Million cell updates/sec

Title: US-09-549-186B-8
Perfect score: 101
Sequence: 1 CGSGVRGDFGSLALRVARQL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	85	84.2	2332	1	GNNY4F	genome polyprotein
2	67	66.3	216	2	A03911	genome polyprotein
3	50	49.5	230	2	A03909	genome polyprotein
4	50	49.5	2336	2	S37077	genome polyprotein
5	48	47.5	564	2	F90965	hypothetical prote
6	48	47.5	564	2	F85813	hypothetical prote
7	48	47.5	569	2	H64959	probable membrane
8	48	47.5	2333	1	GNNY2F	genome polyprotein
9	47	46.5	170	2	T35564	ribosomal protein
10	45	44.6	109	2	S61252	genome polyprotein
11	45	44.6	109	2	S61253	genome polyprotein
12	44	43.6	76	2	D82844	carbon storage reg
13	44	43.6	303	2	E72463	hypothetical prote
14	44	43.6	505	2	S44647	f42h10.1 protein -
15	44	43.6	512	2	S23355	alpha-amylase (EC
16	44	43.6	537	2	AB1869	type I site-specif
17	43.5	43.1	429	2	S45459	TOM34 protein - ye
18	43	42.6	39	2	I49418	insulin I precurso
19	43	42.6	108	1	INMS1	insulin 1 precurso
20	43	42.6	260	2	B64563	hypothetical prote
21	43	42.6	281	1	S34626	translation elonga
22	43	42.6	312	2	A56911	TRADD protein - hu
23	43	42.6	400	2	S20941	protochlorophyllid
24	43	42.6	681	2	E82812	outer membrane hem
25	43	42.6	1159	2	AH3088	IcmF [imported] -
26	43	42.6	1159	2	B98198	hypothetical prote
27	43	42.6	2205	1	GNNY2W	genome polyprotein
28	43	42.6	2207	1	GNNY5P	genome polyprotein
29	43	42.6	2207	2	S09553	genome polyprotein
30	42	41.6	110	1	IPRT1	insulin 1 precurso
31	42	41.6	173	2	T11545	NADH2 dehydrogenas
32	42	41.6	173	2	T11311	NADH2 dehydrogenas
33	42	41.6	291	2	T37992	probable tricarbox
34	42	41.6	309	2	H71328	probable flagellar
35	42	41.6	331	2	E82389	probable outer mem
36	42	41.6	350	1	S52153	alcohol dehydrogen
37	42	41.6	467	2	B75322	probable oligoendo
38	42	41.6	486	1	KRXL	keratin 3, type I,
39	42	41.6	1381	2	S55619	capsid protein 25
40	41.5	41.1	110	1	IPRT2	insulin 2 precurso
41	41.5	41.1	570	2	B86827	hypothetical prote
42	41	40.6	136	2	H90573	ribosomal protein
43	41	40.6	229	2	AD2462	hypothetical prote
44	41	40.6	236	2	E69479	conserved hypothet
45	41	40.6	289	2	A87534	carboxylesterase f

ALIGNMENTS

RESULT 1
GNNY4F
genome polyprotein - foot-and-mouth disease virus A (strain A12)
N;Contains: coat protein VP1; coat protein VP2; coat protein VP3; coat protein VP4; c
tein VPg2; genome-linked protein VPg3; nonstructural protein p20a; proteinase (EC 3.4
C;Species: Aphthovirus A (foot-and-mouth disease virus A)
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C;Accession: A25794
R;Robertson, B.H.; Grubman, M.J.; Weddell, G.N.; Moore, D.M.; Welsh, J.D.; Fischer, T
J. Virol. 54, 651-660, 1985
A;Title: Nucleotide and amino acid sequence coding for polypeptides of foot-and-mouth
A;Reference number: A25794; MUID:85211015; PMID:2987518
A;Accession: A25794
A;Molecule type: genomic RNA
A;Residues: 1-2332 <ROB>
A;Cross-references: GB:M10975; NID:g210306; PIDN:AAA42593.1; PID:g210307
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: coat protein; core protein; genome-linked protein; hydrolase; nonstructur
F;1-216/Product: nonstructural protein p20a #status predicted <NPA>
F;217-285/Product: coat protein VP4 #status predicted <VP4>
F;286-503/Product: coat protein VP2 #status predicted <VP2>
F;504-723/Product: coat protein VP3 #status predicted <VP3>
F;724-937/Product: coat protein VP1 #status predicted <VP1>
F;938-953/Product: core protein X #status predicted <CPX>
F;954-1107/Product: core protein p14 #status predicted <C14>
F;1108-1425/Product: core protein p41 #status predicted <C41>
F;1426-1578/Product: core protein p19 #status predicted <C19>
F;1579-1601/Product: genome-linked protein VPg1 #status predicted <VG1>
F;1602-1625/Product: genome-linked protein VPg2 #status predicted <VG2>
F;1626-1649/Product: genome-linked protein VPg3 #status predicted <VG3>
F;1650-1862/Product: proteinase #status predicted <PTS>
F;1863-2332/Product: RNA-directed RNA polymerase #status predicted <RRP>

Query Match 84.2%; Score 85; DB 1; Length 2332;
Best Local Similarity 94.7%; Pred. No. 4.7e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 GSGVRGDFGSLALRVARQL 20
              ||||||||||||| ||||||
Db        864 GSGVRGDFGSLAPRVARQL 882
```

RESULT 2
A03911
genome polyprotein - foot-and-mouth disease virus A (strain A24 Cruzeiro) (fragment)
N;Contains: coat protein VP1; core protein p52
C;Species: Aphthovirus A (foot-and-mouth disease virus A)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 26-Aug-1999
C;Accession: A03911
R;Makoff, A.J.; Paynter, C.A.; Rowlands, D.J.; Boothroyd, J.C.
Nucleic Acids Res. 10, 8285-8295, 1982
A;Title: Comparison of the amino acid sequence of the major immunogen from three sero

A;Reference number: A37503; MUID:83143292; PMID:6298715
A;Accession: A03911
A;Molecule type: genomic RNA
A;Residues: 1-216 <MAK>
A;Cross-references: GB:J02183; NID:g210312; PIDN:AAA42596.1; PID:g210313
C;Comment: Specific enzymatic cleavage in vivo yields mature proteins including coat pro
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: polyprotein

Query Match 66.3%; Score 67; DB 2; Length 216;
Best Local Similarity 73.7%; Pred. No. 0.0034;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GSGVRGDFGSLALRVARQL 20
||| ||| |||| || :||
Db 129 GSGRRGDMGSLAARVVKQL 147

RESULT 3
A03909
genome polyprotein - foot-and-mouth disease virus A (strain A5) (fragment)
N;Contains: coat protein VP1; coat protein VP3; core protein p52
C;Species: Aphthovirus A (foot-and-mouth disease virus A)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 07-Feb-1997
C;Accession: A03909
R;Beck, E.; Feil, G.; Strohmaier, K.
EMBO J. 2, 555-559, 1983
A;Title: The molecular basis of the antigenic variation of foot-and-mouth disease virus.
A;Reference number: A03909; MUID:84028562; PMID:6194987
A;Accession: A03909
A;Molecule type: mRNA
A;Residues: 1-230 <BEC>
A;Note: the authors translated the codon GCA for residue 114 as Val
C;Comment: Specific enzymatic cleavages in vivo yield mature proteins including the carb
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: polyprotein

Query Match 49.5%; Score 50; DB 2; Length 230;
Best Local Similarity 57.9%; Pred. No. 1.9;
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 2 GSGVRGDFGSLALRVARQL 20
| ||| || | | |:||
Db 143 GGPRRGDMGSAAARAAKQL 161

RESULT 4
S37077
genome polyprotein - foot-and-mouth disease virus A (strain A22/550 Azerbaijan 65)
N;Contains: coat protein VP1; coat protein VP2; coat protein VP3; coat protein VP4; core
tein VPg2; genome-linked protein VPg3; nonstructural protein p20a; proteinase (EC 3.4.-.
C;Species: Aphthovirus A (foot-and-mouth disease virus A)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 26-Aug-1999
C;Accession: S37077; JN0413
R;Sosnovtsev, S.V.; Onischenko, A.M.; Petrov, N.A.; Kalashnikova, T.I.; Mamaeva, N.V.; D
submitted to the EMBL Data Library, August 1993
A;Reference number: S37077
A;Accession: S37077
A;Molecule type: genomic RNA
A;Residues: 1-2336 <SOS>
A;Cross-references: EMBL:X74812; NID:g397965; PIDN:CAA52812.1; PID:g397966
R;Onishchenko, A.M.; Petrov, N.A.; Blinov, V.M.; Vassilenko, S.K.; Sandakhchiev, L.S.; E
Bioorg. Khim. 12, 416-419, 1986
A;Title: The DNA-copy primary structure for the gene of the FMDV A22 serotype VP1-protei
A;Reference number: JN0413; MUID:86186952; PMID:2421736
A;Accession: JN0413
A;Molecule type: genomic RNA
A;Residues: 702-955 <ONI>
A;Cross-references: GB:M38362; NID:g210514; PIDN:AAA42664.1; PID:g210515
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: coat protein; core protein; genome-linked protein; hydrolase; nonstructural
F;1-217/Product: nonstructural protein p20a #status predicted <NPA>
F;218-286/Product: coat protein VP4 #status predicted <VP4>
F;287-504/Product: coat protein VP2 #status predicted <VP2>
F;505-724/Product: coat protein VP3 #status predicted <VP3>
F;725-938/Product: coat protein VP1 #status predicted <VP1>
F;939-954/Product: core protein X #status predicted <CPX>
F;955-1108/Product: core protein p14 #status predicted <C14>
F;1109-1426/Product: core protein p41 #status predicted <C41>
F;1427-1579/Product: core protein p19 #status predicted <C19>
F;1580-1602/Product: genome-linked protein VPg1 #status predicted <VG1>
F;1603-1626/Product: genome-linked protein VPg2 #status predicted <VG2>
F;1627-1650/Product: genome-linked protein VPg3 #status predicted <VG3>
F;1651-1863/Product: proteinase #status predicted <PTS>
F;1864-2333/Product: RNA-directed RNA polymerase #status predicted <RRP>

Query Match 49.5%; Score 50; DB 2; Length 2336;
Best Local Similarity 63.2%; Pred. No. 20;
Matches 12; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 GSGVRGDFGSLALRVARQL 20
| | ||| || ||| ||
Db 865 GMGRRGDLEPLAARVAAQL 883

RESULT 5
F90965
hypothetical protein ECs2694 [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C;Accession: F90965
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90965
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-564 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB36117.1; PID:g13362162; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs2694

Query Match 47.5%; Score 48; DB 2; Length 564;
Best Local Similarity 72.7%; Pred. No. 10;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GVRGDFGSLAL 14
||||||||:::
Db 350 GVRGDFGSISI 360

RESULT 6
F85813
hypothetical protein Z3047 [imported] - Escherichia coli (strain O157:H7, substrain E
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: F85813
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: F85813
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-564 <STO>
A;Cross-references: GB:AE005174; NID:g12516062; PIDN:AAG56970.1; GSPDB:GN00145; UWGP:
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z3047

Query Match 47.5%; Score 48; DB 2; Length 564;
Best Local Similarity 72.7%; Pred. No. 10;

```
  Matches    8;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GVRGDFGSLAL 14
              |||||||:::
Db        350 GVRGDFGSISI 360


RESULT 7
H64959
probable membrane protein b1956 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C;Accession: H64959
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: H64959
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-569 <BLAT>
A;Cross-references: GB:AE000287; GB:U00096; NID:g1788257; PIDN:AAC75022.1; PID:g1788266;
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Start codon: TTG
C;Keywords: transmembrane protein
F;30-46/Domain: transmembrane #status predicted <TM1>
F;363-379/Domain: transmembrane #status predicted <TM2>

  Query Match             47.5%;  Score 48;  DB 2;  Length 569;
  Best Local Similarity   72.7%;  Pred. No. 10;
  Matches    8;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GVRGDFGSLAL 14
              |||||||:::
Db        355 GVRGDFGSISI 365


RESULT 8
GNNY2F
genome polyprotein - foot-and-mouth disease virus A (strain A[10]61)
N;Contains: coat protein VP1; coat protein VP2; coat protein VP3; coat protein VP4; core
 protein p20a; nonstructural protein p20b; RNA-directed RNA polymerase (EC 2.7.7.48)
C;Species: Aphthovirus A (foot-and-mouth disease virus A)
C;Date: 17-Dec-1982 #sequence_revision 28-Aug-1985 #text_change 16-Jul-1999
C;Accession: A93508; A91491; S30753
R;Carroll, A.R.; Rowlands, D.J.; Clarke, B.E.
Nucleic Acids Res. 12, 2461-2472, 1984
A;Title: The complete nucleotide sequence of the RNA coding for the primary translation
A;Reference number: A93508; MUID:84169547; PMID:6324120
A;Accession: A93508
A;Molecule type: genomic RNA
A;Residues: 1-2333 <CAR>
A;Cross-references: GB:X00429
R;Boothroyd, J.C.; Harris, T.J.R.; Rowlands, D.J.; Lowe, P.A.
Gene 17, 153-161, 1982
A;Title: The nucleotide sequence of cDNA coding for the structural proteins of foot-and-
A;Reference number: A91491; MUID:82211814; PMID:6282711
A;Accession: A91491
A;Molecule type: genomic RNA
A;Residues: 115-395,'C',397-631,'L',633-1048 <BOO>
A;Cross-references: GB:V01130; NID:g61048; PIDN:CAA24361.1; PID:g1335402
R;Sangar, D.V.; Newton, S.E.; Rowlands, D.J.; Clarke, B.E.
Nucleic Acids Res. 15, 3305-3315, 1987
A;Title: All foot and mouth disease virus serotypes initiate protein synthesis at two se
A;Reference number: S30753; MUID:87203363; PMID:3033601
A;Accession: S30753
A;Molecule type: genomic RNA
A;Residues: 1-32 <SAN>
A;Cross-references: EMBL:M31575; NID:g210486; PIDN:AAA42655.1; PID:g210487
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: coat protein; core protein; genome-linked protein; nonstructural protein; nu
F;1-204/Product: nonstructural protein p20a #status predicted <NPA>
F;205-286/Product: coat protein VP4 #status predicted <VP4>
F;287-504/Product: coat protein VP2 #status predicted <VP2>
F;505-725/Product: coat protein VP3 #status predicted <VP3>
F;726-937/Product: coat protein VP1 #status predicted <VP1>
F;938-1578/Product: core protein p52 #status predicted <CPP>
F;1579-1601/Product: genome-linked protein VPg1 #status predicted <GL1>
F;1602-1625/Product: genome-linked protein VPg2 #status predicted <GL2>
F;1626-1649/Product: genome-linked protein VPg3 #status predicted <GL3>
F;1650-1863/Product: nonstructural protein p20b #status predicted <NPB>
F;1864-2333/Product: RNA-directed RNA polymerase #status predicted <RRP>

  Query Match             47.5%;  Score 48;  DB 1;  Length 2333;
  Best Local Similarity   71.4%;  Pred. No. 41;
  Matches   10;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 GDFGSLALRVARQL 20
              || ||:| ||| ||
Db        869 GDLGSIAARVATQL 882


RESULT 9
T35564
ribosomal protein S9 - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 19-May-2000
C;Accession: T35564
R;Saunders, D.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1998
A;Reference number: Z21583
A;Accession: T35564
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-170 <SAU>
A;Cross-references: EMBL:AL031317; PIDN:CAA20391.1; GSPDB:GN00070; SCOEDB:SC6G4.13
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: rpsI; SCOEDB:SC6G4.13
C;Superfamily: Escherichia coli ribosomal protein S9

  Query Match             46.5%;  Score 47;  DB 2;  Length 170;
  Best Local Similarity   57.9%;  Pred. No. 4.4;
  Matches   11;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              | || |  |:| | ||| |
Db        108 GGGVSGQAGALRLGVARAL 126


RESULT 10
S61252
genome polyprotein - foot-and-mouth disease virus Asia (isolate Asia I Mandya, Karnat
N;Alternate names: immunogenic polyprotein
C;Species: Aphthovirus Asia (foot-and-mouth disease virus Asia)
A;Variety: isolate Asia I Mandya, Karnataka
C;Date: 18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change 26-Aug-1999
C;Accession: S61252
R;Tulasiram, P.; Tyagi, M.; Suryanarayana, V.
submitted to the EMBL Data Library, June 1995
A;Description: Antigenic variation in foot and mouth disease virus type Asia I isolat
A;Reference number: S61252
A;Accession: S61252
A;Molecule type: mRNA
A;Residues: 1-109 <TUL>
A;Cross-references: EMBL:X88856; NID:g971407; PIDN:CAA61326.1; PID:g971408
A;Experimental source: isolate Asia I Mandya, Karnataka
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: coat protein; polyprotein; proteinase
F;1-91/Product: coat protein VP1 (fragment) #status predicted <MAT1>
F;92-107/Product: proteinase 2A #status predicted <MAT2>
F;108-109/Product: proteinase 2B (fragment) #status predicted <MAT3>

  Query Match             44.6%;  Score 45;  DB 2;  Length 109;
```

```
  Best Local Similarity   60.0%;  Pred. No. 5.9;
  Matches    9;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 RGDFGSLALRVARQL 20
              |||  :|| |::|||
Db         23 RGDMAALAQRLSRQL 37


RESULT 11
S61253
genome polyprotein - foot-and-mouth disease virus Asia (isolate Asia I Nilgiri, Tamil Na
N;Alternate names: immunogenic polyprotein
N;Contains: coat protein VP1; proteinase 2A; proteinase 2B
C;Species: Aphthovirus Asia (foot-and-mouth disease virus Asia)
A;Variety: isolate Asia I Nilgiri, Tamil Nadu
C;Date: 18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change 26-Aug-1999
C;Accession: S61253
R;Tulasiram, P.; Tyagi, M.; Suryanarayana, V.
submitted to the EMBL Data Library, June 1995
A;Description: Antigenic variation in foot and mouth disease virus type Asia I isolates
A;Reference number: S61253
A;Accession: S61253
A;Molecule type: mRNA
A;Residues: 1-109 <TUL>
A;Cross-references: EMBL:X88857; NID:g971409; PIDN:CAA61327.1; PID:g971410
A;Experimental source: isolate Asia I Nilgiri, Tamil Nadu
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: coat protein; polyprotein; proteinase
F;1-91/Product: coat protein VP1 (fragment) #status predicted <MAT1>
F;92-107/Product: proteinase 2A #status predicted <MAT2>
F;108-109/Product: proteinase 2B (fragment) #status predicted <MAT3>

  Query Match            44.6%;  Score 45;  DB 2;  Length 109;
  Best Local Similarity   60.0%;  Pred. No. 5.9;
  Matches    9;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 RGDFGSLALRVARQL 20
              |||  :|| |::|||
Db         23 RGDMAALAQRLSRQL 37


RESULT 12
D82844
carbon storage regulator XF0125 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 17-Nov-2000
C;Accession: D82844
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: D82844
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-76 <SIM>
A;Cross-references: GB:AE003866; GB:AE003849; NID:g9104906; PIDN:AAF82938.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0125
C;Superfamily: glycogen biosynthesis inhibitor

  Query Match            43.6%;  Score 44;  DB 2;  Length 76;
  Best Local Similarity   72.7%;  Pred. No. 5.9;
  Matches    8;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDFGS 11
              | || ||| ||
Db         64 CSSGTRGDSGS 74


RESULT 13
E72463
hypothetical protein APE2352 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C;Accession: E72463
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: E72463
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-303 <KAW>
A;Cross-references: DDBJ:AP000064; NID:g5105945; PIDN:BAA81365.1; PID:d1045151; PID:g
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE2352

  Query Match            43.6%;  Score 44;  DB 2;  Length 303;
  Best Local Similarity   55.6%;  Pred. No. 24;
  Matches   10;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDFGSLALRVAR 18
              ||  | |   ||| |:||
Db        237 CGGPVEGLMPSLAKRIAR 254


RESULT 14
S44647
f42h10.1 protein - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Mar-2001
C;Accession: S44647
R;Du, Z.; Waterston, R.
submitted to the EMBL Data Library, May 1993
A;Description: Sequence of the C. elegans cosmid F42H10.
A;Reference number: S44648
A;Accession: S44647
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-505 <DUZ>
A;Cross-references: EMBL:L08403; NID:g289674; PID:g289675
C;Genetics:
A;Introns: 84/1; 165/3; 215/1; 280/2; 364/3; 414/1; 476/2

  Query Match            43.6%;  Score 44;  DB 2;  Length 505;
  Best Local Similarity   64.7%;  Pred. No. 39;
  Matches   11;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVAR 18
              |  | | ||| ||| ||
Db        134 GDIVDGAFGSWALRAAR 150


RESULT 15
S23355
alpha-amylase (EC 3.2.1.1) precursor - yeast (Schwanniomyces occidentalis)
C;Species: Schwanniomyces occidentalis
C;Date: 22-Nov-1993 #sequence_revision 13-Mar-1997 #text_change 22-Jun-1999
```

C;Accession: S23355
R;Wu, F.M.; Wang, T.T.; Hsu, W.H.
FEMS Microbiol. Lett. 82, 313-318, 1991
A;Title: The nucleotide sequence of Schwanniomyces occidentalis alpha-amylase gene.
A;Reference number: S23355
A;Accession: S23355
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-512 <WUF>
A;Cross-references: EMBL:X62079; NID:g4880; PIDN:CAA43995.1; PID:g4881
A;Note: the authors translated the codon AGA for residue 21 as Pro, CTT for residue 61 a
C;Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C;Keywords: glycoprotein; glycosidase; hydrolase
F;209-336/Domain: alpha-amylase core homology <AMY>

Query Match 43.6%; Score 44; DB 2; Length 512;
Best Local Similarity 47.1%; Pred. No. 40;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 CGSGVRGDFGSLALRVA 17
              | |   ||||||:: ::
Db        475 CNSVSAGDFGSLSVSIS 491
```

Search completed: October 9, 2003, 10:18:42
Job time : 21.6667 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:02:54 ; Search time 10.3333 Seconds
(without alignments)
91.019 Million cell updates/sec

Title: US-09-549-186B-8
Perfect score: 101
Sequence: 1 CGSGVRGDFGSLALRVARQL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	85	84.2	2332	1	POLG_FMDVA	P03308 f genome po
2	67	66.3	216	1	POLG_FMDVC	P03309 foot-and-mo
3	50	49.5	230	1	POLG_FMDV5	P03307 foot-and-mo
4	50	49.5	2336	1	POLG_FMDVZ	P49303 f genome po
5	48	47.5	564	1	YEDQ_ECO57	Q8xb92 escherichia
6	48	47.5	564	1	YEDQ_ECOLI	P76330 escherichia
7	48	47.5	2333	1	POLG_FMDV1	P03306 f genome po
8	47	46.5	170	1	RS9_STRCO	Q53875 streptomyce
9	45	44.6	110	1	INS_PSAOB	Q62587 psammomys o
10	45	44.6	348	1	ADH1_PICST	O00097 pichia stip
11	44	43.6	309	1	YLZ9_CAEEL	P34414 caenorhabdi
12	43.5	43.1	429	1	HRB1_YEAST	P38922 saccharomyc
13	43	42.6	108	1	INS1_MOUSE	P01325 mus musculu
14	43	42.6	129	1	RS9_CHLTE	Q8kbk5 chlorobium
15	43	42.6	281	1	EF1D_HUMAN	P29692 homo sapien
16	43	42.6	312	1	TRAD_HUMAN	Q15628 homo sapien
17	43	42.6	2205	1	POLG_POL2W	P23069 poliovirus
18	43	42.6	2207	1	POLG_POL2L	P06210 poliovirus
19	42	41.6	110	1	INS1_RAT	P01322 rattus norv
20	42	41.6	173	1	NU6M_SCYCA	O79412 scyliorhinu
21	42	41.6	173	1	NU6M_SQUAC	Q9zz43 squalus aca
22	42	41.6	309	1	FLIH_TREPA	O83416 treponema p
23	42	41.6	350	1	ADH1_CANAL	P43067 candida alb
24	42	41.6	486	1	K1C0_XENLA	P02537 xenopus lae
25	41.5	41.1	110	1	INS2_RAT	P01323 rattus norv
26	41	40.6	128	1	RR9_CYACA	Q9tlv4 cyanidium c
27	41	40.6	136	1	RS9_MYCPU	Q98q72 mycoplasma
28	41	40.6	348	1	ADH2_PICST	O13309 pichia stip
29	41	40.6	351	1	NOV_RAT	Q9qzq5 rattus norv
30	41	40.6	351	1	RPOA_TREPA	O83242 treponema p
31	41	40.6	357	1	YNT1_YEAST	P53871 saccharomyc
32	41	40.6	759	1	CIQ2_MOUSE	Q9z351 mus musculu
33	41	40.6	5147	1	FAT_DROME	P33450 drosophila
34	40.5	40.1	555	1	YKQC_BACSU	Q45493 bacillus su
35	40.5	40.1	933	1	PERT_CANFA	Q8hyb7 canis famil
36	40	39.6	130	1	RS9_XANAC	Q8pq41 xanthomonas
37	40	39.6	130	1	RS9_XANCP	Q8pd66 xanthomonas
38	40	39.6	130	1	RS9_XYLFA	Q9pd43 xylella fas
39	40	39.6	134	1	RS9_THEMA	Q9x1g4 thermotoga
40	40	39.6	151	1	RS9_MYCTU	O06259 mycobacteri
41	40	39.6	153	1	RS9_MYCLE	P40828 mycobacteri
42	40	39.6	281	1	WS22_MOUSE	Q9cy21 mus musculu
43	40	39.6	311	1	PUR2_VIGUN	P52421 vigna ungui
44	40	39.6	347	1	ADH1_YEAST	P00330 saccharomyc
45	40	39.6	347	1	ADH2_KLUMA	Q9p4c2 kluyveromyc

ALIGNMENTS

RESULT 1
POLG_FMDVA
ID POLG_FMDVA STANDARD; PRT; 2332 AA.
AC P03308; P03312; Q65038; Q65039; Q65040; Q65041; Q65042; Q65043;
AC Q65044; Q65045; Q65046; Q65047;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat
DE proteins VP1 TO VP4; Core proteins X, P14, P41, P19; Genome-linked
DE proteins VPG1 TO VPG3; Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48)].
OS Foot-and-mouth disease virus (strain A12) (Aphthovirus A) (FMDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Aphthovirus.
OX NCBI_TaxID=12114;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85211015; PubMed=2987518;
RA Robertson B.H., Grubman M.J., Weddell G.N., Moore D.M., Welsh J.D.,
RA Fischer T., Dowbenko D.J., Yansura D.G., Small B., Kleid D.G.;
RT "Nucleotide and amino acid sequence coding for polypeptides of
RT foot-and-mouth disease virus type A12.";
RL J. Virol. 54:651-660(1985).
RN [2]
RP SEQUENCE OF 1863-2332 FROM N.A.
RX MEDLINE=83225613; PubMed=6305004;
RA Robertson B.H., Morgan D.O., Moore D.M., Grubman M.J., Card J.,
RA Fischer T., Weddell G.N., Dowbenko D.J., Yansura D.G.;
RT "Identification of amino acid and nucleotide sequence of the
RT foot-and-mouth disease virus RNA polymerase.";
RL Virology 126:614-623(1983).
RN [3]
RP SEQUENCE OF 715-955 FROM N.A.
RX MEDLINE=82061853; PubMed=6272395;
RA Kleid D.G., Yansura D.G., Small B., Dowbenko D.J., Moore D.M.,
RA Grubman M.J., McKercher P.D., Morgan D.O., Robertson B.H.,
RA Bachrach H.L.;
RT "Cloned viral protein vaccine for foot-and-mouth disease: responses
RT in cattle and swine.";
RL Science 214:1125-1129(1981).
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC --

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M10975; AAA42593.1; -.
DR   EMBL; J02187; AAA42670.1; -.
DR   MEROPS; C03.008; -.
DR   InterPro; IPR004004; Calici_pol_hel.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   InterPro; IPR000605; RNA_helicase.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR001205; RNA_pol_P3D.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00073; rhv; 3.
DR   Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR   Pfam; PF00910; RNA_helicase; 1.
DR   PRINTS; PR00918; CALICVIRUSNS.
DR   PRINTS; PR01542; FMDVP1COAT.
KW   Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;
KW   Transferase; Hydrolase; Thiol protease; Nonstructural protein;
KW   Myristate.
FT   CHAIN         1    200       NONSTRUCTURAL PROTEIN P20A.
FT   CHAIN       201    285       COAT PROTEIN VP4.
FT   CHAIN       286    503       COAT PROTEIN VP2.
FT   CHAIN       504    723       COAT PROTEIN VP3.
FT   CHAIN       724    937       COAT PROTEIN VP1.
FT   CHAIN       938    953       CORE PROTEIN X.
FT   CHAIN       954   1107       CORE PROTEIN P14.
FT   CHAIN      1108   1425       CORE PROTEIN P41.
FT   CHAIN      1426   1578       CORE PROTEIN P19.
FT   CHAIN      1579   1601       GENOME-LINKED PROTEIN VPG1.
FT   CHAIN      1602   1625       GENOME-LINKED PROTEIN VPG2.
FT   CHAIN      1626   1649       GENOME-LINKED PROTEIN VPG3.
FT   CHAIN      1650   1862       PROTEASE.
FT   CHAIN      1863   2332       RNA-DIRECTED RNA POLYMERASE.
FT   LIPID       201    201       MYRISTATE.
SQ   SEQUENCE   2332 AA;  259408 MW;  EE77DA739CBEDC6A CRC64;

  Query Match             84.2%;  Score 85;  DB 1;  Length 2332;
  Best Local Similarity   94.7%;  Pred. No. 2.3e-05;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              |||||||||||| ||||||
Db        864 GSGVRGDFGSLAPRVARQL 882


RESULT 2
POLG_FMDVC
ID   POLG_FMDVC     STANDARD;      PRT;   216 AA.
AC   P03309;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Genome polyprotein [Contains: Coat protein VP1; Core protein P52]
DE   (Fragment).
OS   Foot-and-mouth disease virus (strain A24 Cruzeiro) (Aphthovirus A)
OS   (FMDV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12115;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83143292; PubMed=6298715;
RA   Makoff A.J., Paynter C.A., Rowlands D.J., Boothroyd J.C.;
RT   "Comparison of the amino acid sequence of the major immunogen from
RT   three serotypes of foot and mouth disease virus.";
RL   Nucleic Acids Res. 10:8285-8295(1982).
CC   -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC       EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC       VP3, AND VP4.
CC   -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; J02183; AAA42596.1; -.
DR   EMBL; A06733; CAA00589.1; -.
DR   PIR; A03911; A03911.
DR   HSSP; Q88571; 1TME.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
KW   Coat protein; Core protein; Polyprotein.
FT   NON_TER       1      1
FT   CHAIN         1    202       COAT PROTEIN VP1.
FT   CHAIN       203   >216       CORE PROTEIN P52.
FT   NON_TER     216    216
SQ   SEQUENCE   216 AA;  23889 MW;  501659FF031A1D85 CRC64;

  Query Match             66.3%;  Score 67;  DB 1;  Length 216;
  Best Local Similarity   73.7%;  Pred. No. 0.0015;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              ||| ||| |||| || :||
Db        129 GSGRRGDMGSLAARVVKQL 147


RESULT 3
POLG_FMDV5
ID   POLG_FMDV5     STANDARD;      PRT;   230 AA.
AC   P03307;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Genome polyprotein [Contains: Coat proteins VP3, VP1; Core protein
DE   P52] (Fragment).
OS   Foot-and-mouth disease virus (strain A5) (Aphthovirus A) (FMDV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12113;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84028562; PubMed=6194987;
RA   Beck E., Feil G., Strohmaier K.;
RT   "The molecular basis of the antigenic variation of foot-and-mouth
RT   disease virus.";
RL   EMBO J. 2:555-559(1983).
CC   -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC       EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC       VP3, AND VP4.
CC   -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; V01135; CAA24365.1; ALT_INIT.
DR   EMBL; V01135; CAA24366.1; ALT_SEQ.
DR   PIR; A03909; A03909.
DR   HSSP; Q88571; 1TME.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
```

```
DR   PRINTS; PR01542; FMDVP1COAT.
KW   Coat protein; Core protein; Polyprotein.
FT   NON_TER       1      1
FT   CHAIN        <1      4       COAT PROTEIN VP3.
FT   CHAIN         5    216       COAT PROTEIN VP1.
FT   CHAIN       217   >230       CORE PROTEIN P52.
FT   NON_TER     230    230
SQ   SEQUENCE   230 AA;  25369 MW;  EA93A190F4CC1608 CRC64;

  Query Match             49.5%;  Score 50;  DB 1;  Length 230;
  Best Local Similarity   57.9%;  Pred. No. 0.86;
  Matches   11;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              |   ||| || | | |:||
Db        143 GGPRRGDMGSAAARAAKQL 161


RESULT 4
POLG_FMDVZ
ID   POLG_FMDVZ     STANDARD;      PRT;  2336 AA.
AC   P49303;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Genome polyprotein [Contains: Nonstructural protein P20A; Coat
DE   proteins VP1 TO VP4; Core proteins X, P14, P41, P19; Genome-linked
DE   proteins VPG1 TO VPG3; Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE   (P3C); RNA-directed RNA polymerase (EC 2.7.7.48)].
OS   Foot-and-mouth disease virus (strain A22/550 Azerbaijan 65)
OS   (Aphthovirus A) (FMDV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=73481;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sosnovtsev S.V., Onischenko A.M., Petrov N.A., Kalashnikova T.I.,
RA   Mamaeva N.V., Drygin V.Y., Perevozchikova N.A., Vasilenko S.K.;
RL   Submitted (AUG-1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the
CC       poliovirus polyprotein. In other picornavirus reactions Glu may be
CC       substituted for Gln, and Ser or Thr for Gly.
CC   -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC       {RNA}(N).
CC   -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC       EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC       VP3, AND VP4.
CC   -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC   -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X74812; CAA52812.1; -.
DR   PIR; S37077; S37077.
DR   HSSP; Q88571; 1TME.
DR   MEROPS; C03.008; -.
DR   MEROPS; C28.001; -.
DR   InterPro; IPR004004; Calici_pol_hel.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   InterPro; IPR000605; RNA_helicase.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR001205; RNA_pol_P3D.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00073; rhv; 3.
DR   Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR   Pfam; PF00910; RNA_helicase; 1.
DR   PRINTS; PR00918; CALICVIRUSNS.
DR   PRINTS; PR01542; FMDVP1COAT.
KW   Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;
KW   Transferase; Hydrolase; Thiol protease; Nonstructural protein;
KW   Myristate.
FT   CHAIN         1    201       NONSTRUCTURAL PROTEIN P20A.
FT   CHAIN       202    286       COAT PROTEIN VP4.
FT   CHAIN       287    504       COAT PROTEIN VP2.
FT   CHAIN       505    724       COAT PROTEIN VP3.
FT   CHAIN       725    938       COAT PROTEIN VP1.
FT   CHAIN       939    954       CORE PROTEIN X.
FT   CHAIN       955   1108       CORE PROTEIN P14.
FT   CHAIN      1109   1426       CORE PROTEIN P41.
FT   CHAIN      1427   1579       CORE PROTEIN P19.
FT   CHAIN      1580   1602       GENOME-LINKED PROTEIN VPG1.
FT   CHAIN      1603   1626       GENOME-LINKED PROTEIN VPG2.
FT   CHAIN      1627   1650       GENOME-LINKED PROTEIN VPG3.
FT   CHAIN      1651   1863       PROTEASE.
FT   CHAIN      1864   2336       RNA-DIRECTED RNA POLYMERASE.
FT   LIPID       202    202       MYRISTATE (BY SIMILARITY).
SQ   SEQUENCE   2336 AA;  259982 MW;  15AC2AB022B5B954 CRC64;

  Query Match             49.5%;  Score 50;  DB 1;  Length 2336;
  Best Local Similarity   63.2%;  Pred. No. 9.5;
  Matches   12;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              | | |||   || ||| ||
Db        865 GMGRRGDLEPLAARVAAQL 883


RESULT 5
YEDQ_ECO57
ID   YEDQ_ECO57     STANDARD;      PRT;   564 AA.
AC   Q8XB92;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein yedQ.
GN   YEDQ OR Z3047 OR ECS2694.
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: BELONGS TO THE YAIC / YFIN (E.COLI), YHCK (B.SUBTILIS)
CC       FAMILY.
CC   -!- SIMILARITY: Contains 1 GGDEF domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AE005417; AAG56970.1; -.
DR   EMBL; AP002559; BAB36117.1; -.
DR   PIR; F90965; F90965.
DR   InterPro; IPR000160; GGDEF.
DR   Pfam; PF00990; GGDEF; 1.
DR   SMART; SM00267; DUF1; 1.
DR   TIGRFAMs; TIGR00254; GGDEF; 1.
DR   PROSITE; PS50887; GGDEF; 1.
KW   Hypothetical protein; Transmembrane; Complete proteome.
FT   TRANSMEM     20     40       POTENTIAL.
FT   TRANSMEM    360    380       POTENTIAL.
FT   DOMAIN      428    563       GGDEF.
SQ   SEQUENCE   564 AA;  64287 MW;  BEBC2286ADBAECB0 CRC64;

  Query Match             47.5%;  Score 48;  DB 1;  Length 564;
  Best Local Similarity   72.7%;  Pred. No. 4.6;
  Matches    8;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GVRGDFGSLAL 14
              ||||||||:::
Db        350 GVRGDFGSISI 360


RESULT 6
YEDQ_ECOLI
ID   YEDQ_ECOLI     STANDARD;      PRT;   564 AA.
AC   P76330; P94746;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein yedQ.
GN   YEDQ OR B1956.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251358; PubMed=9097040;
RA   Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA   Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
RA   Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
RA   Nashimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA   Sivasundaram S., Tagami H., Takeda J., Takemoto K., Wada C.,
RA   Yamamoto Y., Horiuchi T.;
RT   "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 40.1-50.0 min region on the linkage map.";
RL   DNA Res. 3:379-392(1996).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: BELONGS TO THE YAIC / YFIN (E.COLI), YHCK (B.SUBTILIS)
CC       FAMILY.
CC   -!- SIMILARITY: Contains 1 GGDEF domain.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AE000287; AAC75022.1; ALT_INIT.
DR   EMBL; D90835; BAA15784.1; -.
DR   EcoGene; EG14040; yedQ.
DR   InterPro; IPR000160; GGDEF.
DR   Pfam; PF00990; GGDEF; 1.
DR   SMART; SM00267; DUF1; 1.
DR   TIGRFAMs; TIGR00254; GGDEF; 1.
DR   PROSITE; PS50887; GGDEF; 1.
KW   Hypothetical protein; Transmembrane; Complete proteome.
FT   TRANSMEM     20     40       POTENTIAL.
FT   TRANSMEM    360    380       POTENTIAL.
FT   DOMAIN      428    563       GGDEF.
SQ   SEQUENCE   564 AA;  64283 MW;  05FB02C1BE2A8938 CRC64;

  Query Match             47.5%;  Score 48;  DB 1;  Length 564;
  Best Local Similarity   72.7%;  Pred. No. 4.6;
  Matches    8;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GVRGDFGSLAL 14
              ||||||||:::
Db        350 GVRGDFGSISI 360


RESULT 7
POLG_FMDV1
ID   POLG_FMDV1     STANDARD;      PRT;  2333 AA.
AC   P03306; Q64768; Q84750; Q84751; Q84752; Q84753; Q84754; Q84760;
AC   Q84761; Q84762; Q84763; Q84764; Q84765; Q84766; Q84767; Q84768;
AC   Q84769; Q89824;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Genome polyprotein [Contains: Nonstructural protein P20A; Coat
DE   proteins VP1 TO VP4; Core protein P52; Genome-linked proteins VPG1 TO
DE   VPG3; Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed
DE   RNA polymerase P56A (EC 2.7.7.48)].
OS   Foot-and-mouth disease virus (strain A10-61) (Aphthovirus A) (FMDV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12112;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84169547; PubMed=6324120;
RA   Carroll A.R., Rowlands D.J., Clarke B.E.;
RT   "The complete nucleotide sequence of the RNA coding for the primary
RT   translation product of foot and mouth disease virus.";
RL   Nucleic Acids Res. 12:2461-2472(1984).
RN   [2]
RP   SEQUENCE OF 115-1048 FROM N.A.
RX   MEDLINE=82211814; PubMed=6282711;
RA   Boothroyd J.C., Harris T.J.R., Rowlands D.J., Lowe P.A.;
RT   "The nucleotide sequence of cDNA coding for the structural proteins
RT   of foot-and-mouth disease virus.";
RL   Gene 17:153-161(1982).
CC   -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the
CC       poliovirus polyprotein. In other picornavirus reactions Glu may be
CC       substituted for Gln, and Ser or Thr for Gly.
CC   -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC       {RNA}(N).
CC   -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC       EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC       VP3, AND VP4.
CC   -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC   -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; V01130; CAA24361.1; -.
DR   EMBL; X00429; CAA25127.1; -.
DR   MEROPS; C03.008; -.
DR   MEROPS; C28.001; -.
DR   InterPro; IPR004004; Calici_pol_hel.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   InterPro; IPR000605; RNA_helicase.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR001205; RNA_pol_P3D.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00073; rhv; 3.
DR   Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR   Pfam; PF00910; RNA_helicase; 1.
DR   PRINTS; PR00918; CALICVIRUSNS.
DR   PRINTS; PR01542; FMDVP1COAT.
KW   Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;
KW   Transferase; Hydrolase; Thiol protease; Nonstructural protein;
KW   Myristate.
FT   CHAIN         1    201       NONSTRUCTURAL PROTEIN P20A.
FT   CHAIN       202    286       COAT PROTEIN VP4.
FT   CHAIN       287    504       COAT PROTEIN VP2.
FT   CHAIN       505    725       COAT PROTEIN VP3.
FT   CHAIN       726    937       COAT PROTEIN VP1.
FT   CHAIN       938   1578       CORE PROTEIN P52.
FT   CHAIN      1579   1601       GENOME-LINKED PROTEIN VPG1.
FT   CHAIN      1602   1625       GENOME-LINKED PROTEIN VPG2.
FT   CHAIN      1626   1649       GENOME-LINKED PROTEIN VPG3.
FT   CHAIN      1650   1863       PROTEASE P20B.
FT   CHAIN      1864   2333       RNA-DIRECTED RNA POLYMERASE P56A.
FT   LIPID       202    202       MYRISTATE.
FT   CONFLICT    396    396       S -> C (IN REF. 2).
FT   CONFLICT    632    632       P -> L (IN REF. 2).
SQ   SEQUENCE   2333 AA;  259645 MW;  4FC667DCC521BC60 CRC64;

  Query Match             47.5%;  Score 48;  DB 1;  Length 2333;
  Best Local Similarity   71.4%;  Pred. No. 20;
  Matches   10;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 GDFGSLALRVARQL 20
              || ||:| ||| ||
Db        869 GDLGSIAARVATQL 882


RESULT 8
RS9_STRCO
ID   RS9_STRCO      STANDARD;      PRT;   170 AA.
AC   Q53875;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   30S ribosomal protein S9.
GN   RPSI OR SCO4735 OR SC6G4.13.
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=98102746; PubMed=9439573;
RA   Sanchez C., Blanco G., Mendez C., Salas J.A.;
RT   "Cloning, sequencing and transcriptional analysis of a Streptomyces
RT   coelicolor operon containing the rplM and rpsI genes encoding
RT   ribosomal proteins ScoL13 and ScoS9.";
RL   Mol. Gen. Genet. 257:91-96(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2) / M145;
RX   MEDLINE=21996410; PubMed=12000953;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
CC   -!- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U43429; AAC46061.1; -.
DR   EMBL; AL939121; CAA20391.1; -.
DR   PIR; T35564; T35564.
DR   HAMAP; MF_00532; -; 1.
DR   InterPro; IPR000754; Ribosomal_S9.
DR   Pfam; PF00380; Ribosomal_S9; 1.
DR   ProDom; PD001627; Ribosomal_S9; 1.
DR   PROSITE; PS00360; RIBOSOMAL_S9; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   170 AA;  18699 MW;  3CEE00FF69151C99 CRC64;

  Query Match             46.5%;  Score 47;  DB 1;  Length 170;
  Best Local Similarity   57.9%;  Pred. No. 1.9;
  Matches   11;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              | || |  |:| | ||| |
Db        108 GGGVSGQAGALRLGVARAL 126


RESULT 9
INS_PSAOB
ID   INS_PSAOB      STANDARD;      PRT;   110 AA.
AC   Q62587;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Insulin precursor.
GN   INS.
OS   Psammomys obesus.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
OC   Psammomys.
OX   NCBI_TaxID=48139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreas;
RX   MEDLINE=97309250; PubMed=9166665;
RA   Kaiser N., Bailyes E.M., Schneider B.S., Cerasi E., Steiner D.F.,
RA   Hutton J.C., Gross D.J.;
RT   "Characterization of the unusual insulin of Psammomys obesus, a
RT   rodent with nutrition-induced NIDDM-like syndrome.";
RL   Diabetes 46:953-957(1997).
CC   -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC       INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC       FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
```

```
CC       DISULFIDE BONDS.
CC    -!- SUBCELLULAR LOCATION: Secreted.
CC    -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC    ---------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    ---------------------------------------------------------------------
DR    EMBL; X98241; CAA66897.1; -.
DR    HSSP; P01308; 1AI0.
DR    InterPro; IPR004825; Ins/IGF/relax.
DR    Pfam; PF00049; Insulin; 1.
DR    PRINTS; PR00277; INSULINB.
DR    SMART; SM00078; IlGF; 1.
DR    PROSITE; PS00262; INSULIN; 1.
KW    Insulin family; Hormone; Glucose metabolism; Signal.
FT    SIGNAL        1     24       BY SIMILARITY.
FT    CHAIN        25     54       INSULIN B CHAIN.
FT    PROPEP       57     87       C PEPTIDE.
FT    CHAIN        90    110       INSULIN A CHAIN.
FT    DISULFID     31     96       INTERCHAIN (BY SIMILARITY).
FT    DISULFID     43    109       INTERCHAIN (BY SIMILARITY).
FT    DISULFID     95    100       BY SIMILARITY.
SQ    SEQUENCE   110 AA;  12324 MW;  A006738E20579CB0 CRC64;

  Query Match             44.6%;  Score 45;  DB 1;  Length 110;
  Best Local Similarity   61.1%;  Pred. No. 2.5;
  Matches   11;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQ 19
              ||   ||  :||| ||||
Db         70 GSPGAGDLRALALEVARQ 87


RESULT 10
ADH1_PICST
ID    ADH1_PICST     STANDARD;      PRT;   348 AA.
AC    O00097;
DT    15-DEC-1998 (Rel. 37, Created)
DT    15-DEC-1998 (Rel. 37, Last sequence update)
DT    15-SEP-2003 (Rel. 42, Last annotation update)
DE    Alcohol dehydrogenase I (EC 1.1.1.1) (ADH 2).
GN    ADH1 OR ADH2.
OS    Pichia stipitis (Yeast).
OC    Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC    Saccharomycetales; Saccharomycetaceae; Pichia.
OX    NCBI_TaxID=4924;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=ATCC 58785 / CBS 6054;
RX    MEDLINE=98207839; PubMed=9546172;
RA    Cho J.Y., Jeffries T.W.;
RT    "Pichia stipitis genes for alcohol dehydrogenase with fermentative
RT    and respiratory functions.";
RL    Appl. Environ. Microbiol. 64:1350-1358(1998).
RN    [2]
RP    SEQUENCE FROM N.A.
RC    STRAIN=CBS 5774;
RX    MEDLINE=99019018; PubMed=9802210;
RA    Passoth V., Schaefer B., Liebel B., Weierstall T., Klinner U.;
RT    "Molecular cloning of alcohol dehydrogenase genes of the yeast Pichia
RT    stipitis and identification of the fermentative ADH.";
RL    Yeast 14:1311-1325(1998).
CC    -!- FUNCTION: CONVERTS ETHANOL TO ACETALDEHYDE AND PLAYS A MAJOR ROLE
CC        IN XYLOSE FERMENTATION.
CC    -!- CATALYTIC ACTIVITY: An alcohol + NAD(+) = an aldehyde or ketone +
CC        NADH.
CC    -!- COFACTOR: Binds 2 zinc ions per subunit (By similarity).
CC    -!- SUBUNIT: Homotetramer (By similarity).
CC    -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC    -!- SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase
CC        family.
CC    ---------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    ---------------------------------------------------------------------
DR    EMBL; AF008245; AAC49991.1; -.
DR    EMBL; Y13397; CAA73827.1; -.
DR    InterPro; IPR002328; ADH_zinc.
DR    InterPro; IPR002085; Adh_zn_family.
DR    Pfam; PF00107; ADH_zinc_N; 1.
DR    PROSITE; PS00059; ADH_ZINC; 1.
KW    Oxidoreductase; Zinc; Metal-binding; NAD; Multigene family.
FT    METAL         44     44       ZINC 1 (CATALYTIC) (BY SIMILARITY).
FT    METAL         67     67       ZINC 1 (CATALYTIC) (BY SIMILARITY).
FT    METAL         98     98       ZINC 2 (BY SIMILARITY).
FT    METAL        101    101       ZINC 2 (BY SIMILARITY).
FT    METAL        104    104       ZINC 2 (BY SIMILARITY).
FT    METAL        112    112       ZINC 2 (BY SIMILARITY).
FT    METAL        154    154       ZINC 1 (CATALYTIC) (BY SIMILARITY).
SQ    SEQUENCE   348 AA;  36520 MW;  49C06B545D5350F4 CRC64;

  Query Match             44.6%;  Score 45;  DB 1;  Length 348;
  Best Local Similarity   45.0%;  Pred. No. 8.4;
  Matches    9;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDFGSLALRVARQL 20
              | ||  |  ||||:: |: :
Db        175 CISGAGGGLGSLAIQYAKAM 194


RESULT 11
YLZ9_CAEEL
ID    YLZ9_CAEEL     STANDARD;      PRT;   309 AA.
AC    P34414;
DT    01-FEB-1994 (Rel. 28, Created)
DT    28-FEB-2003 (Rel. 41, Last sequence update)
DT    28-FEB-2003 (Rel. 41, Last annotation update)
DE    Hypothetical protein F42H10.9 in chromosome III.
GN    F42H10.9/F42H10.1.
OS    Caenorhabditis elegans.
OC    Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC    Rhabditidae; Peloderinae; Caenorhabditis.
OX    NCBI_TaxID=6239;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=Bristol N2;
RX    MEDLINE=94150718; PubMed=7906398;
RA    Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA    Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA    Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA    Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA    Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA    Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA    Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA    Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA    Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA    Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA    Wohldman P.;
RT    "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT    elegans.";
RL    Nature 368:32-38(1994).
RN    [2]
RP    REVISIONS.
RA    Waterston R.;
RL    Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
CC    ---------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L08403; AAA28019.2; -.
DR   WormPep; F42H10.9; CE24970.
KW   Hypothetical protein.
SQ   SEQUENCE   309 AA;  35071 MW;  E2D47102CEC52B0A CRC64;

  Query Match             43.6%;  Score 44;  DB 1;  Length 309;
  Best Local Similarity   64.7%;  Pred. No. 11;
  Matches   11;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVAR 18
              |  | | ||| ||| ||
Db        134 GDIVDGAFGSWALRAAR 150


RESULT 12
HRB1_YEAST
ID   HRB1_YEAST     STANDARD;      PRT;   429 AA.
AC   P38922;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   HRB1 protein (TOM34 protein).
GN   HRB1 OR TOM34 OR YNL004W OR N2009.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GRF88;
RX   MEDLINE=94005822; PubMed=8402262;
RA   Lalo D., Stettler S., Mariotte S., Slonimski P.P., Thuriaux P.;
RT   "Two yeast chromosomes are related by a fossil duplication of their
RT   centromeric regions.";
RL   C. R. Acad. Sci., III, Sci. Vie 316:367-373(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GRF88;
RX   MEDLINE=95028151; PubMed=7941739;
RA   Lalo D., Stettler S., Mariotte S., Gendreau E., Thuriaux P.;
RT   "Organization of the centromeric region of chromosome XIV in
RT   Saccharomyces cerevisiae.";
RL   Yeast 10:523-533(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / FY1679;
RX   MEDLINE=95076713; PubMed=7985421;
RA   Verhasselt P., Aert R., Voet M., Volckaert G.;
RT   "Nucleotide sequence analysis of an 8887 bp region of the left arm of
RT   yeast chromosome XIV, encompassing the centromere sequence.";
RL   Yeast 10:945-951(1994).
CC   -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC   -!- SIMILARITY: STRONG, TO YEAST GBP2.
CC   -!- SIMILARITY: Contains 3 RNA recognition motif (RRM) domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U02536; AAA64803.1; -.
DR   EMBL; X77114; CAA54378.1; -.
DR   EMBL; Z71280; CAA95863.1; -.
DR   PIR; S45459; S45459.
DR   HSSP; P09651; 1HA1.
DR   SGD; S0004949; HRB1.
DR   GO; GO:0005737; C:cytoplasm; IDA.
DR   GO; GO:0030530; C:heterogeneous nuclear ribonucleoprotein com. . .; NAS.
DR   GO; GO:0006606; P:protein-nucleus import; IPI.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   Pfam; PF00076; rrm; 3.
DR   SMART; SM00360; RRM; 3.
DR   PROSITE; PS50102; RRM; 3.
DR   PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW   RNA-binding; Nuclear protein; Repeat.
FT   DOMAIN      136    212       RNA-BINDING (RRM) 1.
FT   DOMAIN      236    313       RNA-BINDING (RRM) 2.
FT   DOMAIN      351    428       RNA-BINDING (RRM) 3.
SQ   SEQUENCE   429 AA;  49141 MW;  DC0F732EFA43EE89 CRC64;

  Query Match             43.1%;  Score 43.5;  DB 1;  Length 429;
  Best Local Similarity   55.6%;  Pred. No. 18;
  Matches   10;  Conservative    3;  Mismatches    2;  Indels    3;  Gaps    1;

Qy          3 SGVRGDFGSLALRVARQL 20
              || |||:| |   :||:|
Db        108 SGARGDYGPL---LAREL 122


RESULT 13
INS1_MOUSE
ID   INS1_MOUSE     STANDARD;      PRT;   108 AA.
AC   P01325; Q9D907;
DT   21-JUL-1986 (Rel. 01, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Insulin 1 precursor.
GN   INS1 OR INS-1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87169768; PubMed=3104603;
RA   Wentworth B.M., Schaefer I.M., Villa-Komaroff L., Chirgwin J.M.;
RT   "Characterization of the two nonallelic genes encoding mouse
RT   preproinsulin.";
RL   J. Mol. Evol. 23:305-312(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NON;
RX   MEDLINE=90372989; PubMed=2397023;
RA   Sawa T., Ohgaku S., Morioka H., Yano S.;
RT   "Molecular cloning and DNA sequence analysis of preproinsulin genes
RT   in the NON mouse, an animal model of human non-obese, non-insulin-
RT   dependent diabetes mellitus.";
RL   J. Mol. Endocrinol. 5:61-67(1990).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Pancreas;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
```

```
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [4]
RP   SEQUENCE OF 25-54 AND 88-108.
RX   MEDLINE=72189455; PubMed=5063718;
RA   Buenzli H.F., Glatthaar B., Kunz P., Muelhaupt E., Humbel R.E.;
RT   "Amino acid sequence of the two insulins from mouse (Maus musculus).";
RL   Hoppe-Seyler's Z. Physiol. Chem. 353:451-458(1972).
CC   -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC       INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC       FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC       DISULFIDE BONDS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X04725; CAA28434.1; -.
DR   EMBL; AK007482; BAB25058.1; -.
DR   PIR; B26342; INMS1.
DR   HSSP; P01308; 1A7F.
DR   MGD; MGI:96572; Ins1.
DR   InterPro; IPR004825; Ins/IGF/relax.
DR   Pfam; PF00049; Insulin; 1.
DR   SMART; SM00078; IlGF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Insulin family; Hormone; Glucose metabolism; Signal; Multigene family.
FT   SIGNAL       1     24
FT   CHAIN       25     54       INSULIN 1 B CHAIN.
FT   PROPEP      57     85       INSULIN 1 C PEPTIDE.
FT   CHAIN       88    108       INSULIN 1 A CHAIN.
FT   DISULFID    31     94       INTERCHAIN.
FT   DISULFID    43    107       INTERCHAIN.
FT   DISULFID    93     98
FT   CONFLICT    37     37       E -> K (IN REF. 3).
SQ   SEQUENCE   108 AA;  12160 MW;  F63D9B7B896E0F88 CRC64;

  Query Match             42.6%;  Score 43;  DB 1;  Length 108;
  Best Local Similarity   62.5%;  Pred. No. 5.2;
  Matches   10;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 GVRGDFGSLALRVARQ 19
              |  ||  :||| ||||
Db         70 GSPGDLQTLALEVARQ 85


RESULT 14
RS9_CHLTE
ID   RS9_CHLTE      STANDARD;      PRT;   129 AA.
AC   Q8KBK5;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   30S ribosomal protein S9.
GN   RPSI OR CT1782.
OS   Chlorobium tepidum.
OC   Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC   Chlorobium.
OX   NCBI_TaxID=1097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TLS / ATCC 49652 / DSM 12025;
RX   MEDLINE=22103685; PubMed=12093901;
RA   Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA   Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
RA   Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
RA   Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA   Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
RA   Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA   Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT   "The complete genome sequence of Chlorobium tepidum TLS, a
RT   photosynthetic, anaerobic, green-sulfur bacterium.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
CC   -!- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE012931; AAM73003.1; -.
DR   TIGR; CT1782; -.
DR   HAMAP; MF_00532; -; 1.
DR   InterPro; IPR000754; Ribosomal_S9.
DR   Pfam; PF00380; Ribosomal_S9; 1.
DR   ProDom; PD001627; Ribosomal_S9; 1.
DR   PROSITE; PS00360; RIBOSOMAL_S9; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   129 AA;  14541 MW;  D110858BA87B9BAC CRC64;

  Query Match             42.6%;  Score 43;  DB 1;  Length 129;
  Best Local Similarity   42.1%;  Pred. No. 6.3;
  Matches    8;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              | |: |  |:::| :|| |
Db         68 GGGLTGQSGAVSLAIARAL 86


RESULT 15
EF1D_HUMAN
ID   EF1D_HUMAN     STANDARD;      PRT;   281 AA.
AC   P29692; Q969J1;
DT   01-APR-1993 (Rel. 25, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Elongation factor 1-delta (EF-1-delta) (Antigen NY-CO-4).
GN   EEF1D OR EF1D.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin fibroblast;
RX   MEDLINE=93326642; PubMed=8334168;
RA   Sanders J.P., Raggiaschi R., Morales J., Moeller W.;
RT   "The human leucine zipper-containing guanine-nucleotide exchange
RT   protein elongation factor-1 delta.";
RL   Biochim. Biophys. Acta 1174:87-90(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Colorectal carcinoma;
RX   MEDLINE=98272252; PubMed=9610721;
RA   Scanlan M.J., Chen Y.-T., Williamson B., Gure A.O., Stockert E.,
RA   Gordan J.D., Tureci O., Sahin U., Pfreundschuh M., Old L.J.;
RT   "Characterization of human colon cancer antigens recognized by
RT   autologous antibodies.";
RL   Int. J. Cancer 76:652-658(1998).
RN   [3]
```

```
RP   SEQUENCE FROM N.A.
RC   TISSUE=Muscle;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: EF-1-BETA AND EF-1-DELTA STIMULATE THE EXCHANGE OF
CC       GDP BOUND TO EF-1-ALPHA TO GTP.
CC   -!- SUBUNIT: EF-1 IS COMPOSED OF FOUR SUBUNITS: ALPHA, BETA,
CC       DELTA, AND GAMMA.
CC   -!- SIMILARITY: BELONGS TO THE EF-1-BETA/EF-1-DELTA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z21507; CAA79716.1; -.
DR   EMBL; BC008012; AAH08012.1; -.
DR   EMBL; BC009907; AAH09907.1; -.
DR   EMBL; BC012819; AAH12819.1; -.
DR   PIR; S34626; S34626.
DR   Genew; HGNC:3211; EEF1D.
DR   MIM; 130592; -.
DR   GO; GO:0005853; C:eukaryotic translation elongation factor 1 . . .; TAS.
DR   InterPro; IPR001326; EF1_BD.
DR   Pfam; PF00736; EF1BD; 1.
DR   PROSITE; PS00824; EF1BD_1; 1.
DR   PROSITE; PS00825; EF1BD_2; 1.
KW   Elongation factor; Protein biosynthesis.
FT   CONFLICT     34     34       A -> R (IN REF. 1).
FT   CONFLICT     44     44       S -> T (IN REF. 1).
SQ   SEQUENCE   281 AA;  31122 MW;  CE778D6D5D09BD6C CRC64;

  Query Match             42.6%;  Score 43;  DB 1;  Length 281;
  Best Local Similarity   53.3%;  Pred. No. 14;
  Matches    8;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 SGVRGDFGSLALRVA 17
              ||  || | | :|:|
Db         71 SGTSGDHGELVVRIA 85


Search completed: October  9, 2003, 10:14:58
Job time : 11.3333 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:09:40 ; Search time 49.3333 Seconds
(without alignments)
104.616 Million cell updates/sec

Title: US-09-549-186B-8
Perfect score: 101
Sequence: 1 CGSGVRGDFGSLALRVARQL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	77	76.2	20	12	Q9PXE4	Q9pxe4 foot-and-mo
2	71	70.3	213	12	Q912L2	Q912l2 foot-and-mo
3	71	70.3	213	12	Q67446	Q67446 foot-and-mo
4	71	70.3	213	12	Q67448	Q67448 foot-and-mo
5	71	70.3	213	12	Q65093	Q65093 foot-and-mo
6	68	67.3	213	12	Q98VZ6	Q98vz6 foot-and-mo
7	67	66.3	210	12	Q67438	Q67438 foot-and-mo
8	67	66.3	969	12	Q9Q2N6	Q9q2n6 foot-and-mo
9	66	65.3	169	12	Q67442	Q67442 foot-and-mo
10	65	64.4	126	12	Q8V443	Q8v443 foot-and-mo
11	62	61.4	125	12	Q8V442	Q8v442 foot-and-mo
12	62	61.4	143	12	Q9YPU0	Q9ypu0 foot-and-mo
13	62	61.4	208	12	Q8JUR8	Q8jur8 foot-and-mo
14	62	61.4	211	12	Q8JUP0	Q8jup0 foot-and-mo
15	62	61.4	212	12	Q8JUP1	Q8jup1 foot-and-mo
16	62	61.4	213	12	Q8JUS6	Q8jus6 foot-and-mo
17	62	61.4	213	12	Q8JUM1	Q8jum1 foot-and-mo
18	62	61.4	213	12	Q8JUQ9	Q8juq9 foot-and-mo
19	62	61.4	213	12	Q8JUP7	Q8jup7 foot-and-mo
20	62	61.4	213	12	Q8JUP4	Q8jup4 foot-and-mo
21	62	61.4	213	12	Q912L1	Q912l1 foot-and-mo
22	62	61.4	213	12	Q8JUQ5	Q8juq5 foot-and-mo
23	62	61.4	213	12	Q8JUL1	Q8jul1 foot-and-mo
24	62	61.4	213	12	Q8JUN1	Q8jun1 foot-and-mo
25	62	61.4	213	12	Q8JUP5	Q8jup5 foot-and-mo
26	62	61.4	213	12	Q8JUN0	Q8jun0 foot-and-mo
27	62	61.4	213	12	Q8JUP3	Q8jup3 foot-and-mo
28	62	61.4	213	12	Q8JUQ2	Q8juq2 foot-and-mo
29	62	61.4	213	12	Q8JUQ1	Q8juq1 foot-and-mo
30	62	61.4	213	12	Q8JUL0	Q8jul0 foot-and-mo
31	62	61.4	213	12	Q8JUS7	Q8jus7 foot-and-mo
32	62	61.4	213	12	Q8JUM0	Q8jum0 foot-and-mo
33	62	61.4	213	12	Q8JUM3	Q8jum3 foot-and-mo
34	62	61.4	213	12	Q8JUQ8	Q8juq8 foot-and-mo
35	62	61.4	213	12	Q8JUN9	Q8jun9 foot-and-mo
36	62	61.4	213	12	Q8JUR0	Q8jur0 foot-and-mo
37	62	61.4	213	12	Q8JUS5	Q8jus5 foot-and-mo
38	60	59.4	197	12	Q8JUT2	Q8jut2 foot-and-mo
39	60	59.4	213	12	Q8JUQ4	Q8juq4 foot-and-mo
40	60	59.4	652	12	Q9Q2N8	Q9q2n8 foot-and-mo
41	59	58.4	213	12	Q8JUM9	Q8jum9 foot-and-mo
42	59	58.4	213	12	Q912K6	Q912k6 foot-and-mo
43	59	58.4	213	12	Q66924	Q66924 foot-and-mo
44	59	58.4	213	12	Q912K4	Q912k4 foot-and-mo
45	59	58.4	213	12	Q912L0	Q912l0 foot-and-mo

ALIGNMENTS

RESULT 1
Q9PXE4
ID Q9PXE4 PRELIMINARY; PRT; 20 AA.
AC Q9PXE4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE VP1 protein (Fragment).
OS Foot-and-mouth disease virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Aphthovirus.
OX NCBI_TaxID=12110;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96100820; PubMed=7483796;
RA Piatti P., Hassard S., Newman J.F., Brown F.;
RT "Antigenic variants in a plaque-isolate of foot-and-mouth disease
RT virus: implications for vaccine production.";
RL Vaccine 13:781-784(1995).
SQ SEQUENCE 20 AA; 1925 MW; E91F87C99C19D7DB CRC64;

Query Match 76.2%; Score 77; DB 12; Length 20;
Best Local Similarity 89.5%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GSGVRGDFGSLALRVARQL 20
||||||| ||||||||| |
Db 1 GSGVRGDSGSLALRVARGL 19

RESULT 2
Q912L2
ID Q912L2 PRELIMINARY; PRT; 213 AA.
AC Q912L2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE VP1 protein (Fragment).
GN 1D.

```
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A Arg/68;
RX   MEDLINE=21580796; PubMed=11724271;
RA   Konig G.A., Blanco C., Knowles N.J., Palma E.L., Maradei E.,
RA   Piccone M.E.;
RT   "Phylogenetic analysis of foot-and-mouth disease viruses isolated in
RT   Argentina.";
RL   Virus Genes 23:175-182(2001).
DR   EMBL; AJ308694; CAC48168.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     213    213
SQ   SEQUENCE   213 AA;  23488 MW;  203EFCB8AB45EECE CRC64;

  Query Match             70.3%;  Score 71;  DB 12;  Length 213;
  Best Local Similarity   78.9%;  Pred. No. 0.0017;
  Matches   15;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              ||| ||| |||| |||:||
Db        140 GSGRRGDMGSLAARVAKQL 158


RESULT 3
Q67446
ID   Q67446      PRELIMINARY;      PRT;   213 AA.
AC   Q67446; Q67447;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Foot and mouth disease virus (Serotype Aven76), capsid protein VP1
DE   mRNA (Fragment).
OS   Foot-and-mouth disease virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12110;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85190547; PubMed=2986125;
RA   Weddell G.N., Yansura D.G., Dowbenko D.J., Hoatlin M.E., Grubman M.J.,
RA   Moore D.M., Kleid D.G.;
RT   "Sequence variation in the gene for the immunogenic capsid protein VP1
RT   of foot-and-mouth disease virus type A.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:2618-2622(1985).
DR   EMBL; K03344; AAA42600.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     213    213
SQ   SEQUENCE   213 AA;  23311 MW;  C4ED4E0116A0DB8A CRC64;

  Query Match             70.3%;  Score 71;  DB 12;  Length 213;
  Best Local Similarity   78.9%;  Pred. No. 0.0017;
  Matches   15;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              ||| ||| |||| |||:||
Db        140 GSGRRGDMGSLAARVAKQL 158


RESULT 4
Q67448
ID   Q67448      PRELIMINARY;      PRT;   213 AA.
AC   Q67448; Q67449;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Foot and mouth disease virus (Serotype Aarg79), capsid protein VP1
DE   mRNA (Fragment).
OS   Foot-and-mouth disease virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12110;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85190547; PubMed=2986125;
RA   Weddell G.N., Yansura D.G., Dowbenko D.J., Hoatlin M.E., Grubman M.J.,
RA   Moore D.M., Kleid D.G.;
RT   "Sequence variation in the gene for the immunogenic capsid protein VP1
RT   of foot-and-mouth disease virus type A.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:2618-2622(1985).
DR   EMBL; K03345; AAA42601.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     213    213
SQ   SEQUENCE   213 AA;  23345 MW;  546C7FF78CD45CC0 CRC64;

  Query Match             70.3%;  Score 71;  DB 12;  Length 213;
  Best Local Similarity   78.9%;  Pred. No. 0.0017;
  Matches   15;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              ||| |+| |||| |||:||
Db        140 GSGRRGDMGSLAARVAKQL 158


RESULT 5
Q65093
ID   Q65093      PRELIMINARY;      PRT;   213 AA.
AC   Q65093;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Foot and mouth disease virus strain A Venceslau VP1 (Fragment).
OS   Foot-and-mouth disease virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12110;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85077620; PubMed=6096217;
RA   Cheung A., Whitehead P., Weiss S., Kupper H.;
RT   "Nucleotide sequence of the VP1 gene of the foot-and-mouth disease
RT   virus strain A Venceslau.";
RL   Gene 30:241-245(1984).
DR   EMBL; M12905; AAA42663.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     213    213
SQ   SEQUENCE   213 AA;  23327 MW;  048088D9571D7073 CRC64;

  Query Match             70.3%;  Score 71;  DB 12;  Length 213;
  Best Local Similarity   78.9%;  Pred. No. 0.0017;
  Matches   15;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              ||| ||| |||| |||:||
Db        140 GSGRRGDMGSLAARVAKQL 158
```

```
RESULT 6
Q98VZ6
ID   Q98VZ6      PRELIMINARY;      PRT;   213 AA.
AC   Q98VZ6;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   VP1 (Fragment).
GN   1D.
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A76/Argentina/76;
RX   MEDLINE=21580796; PubMed=11724271;
RA   Konig G.A., Blanco C., Knowles N.J., Palma E.L., Maradei E.,
RA   Piccone M.E.;
RT   "Phylogenetic analysis of foot-and-mouth disease viruses isolated in
RT   Argentina.";
RL   Virus Genes 23:175-182(2001).
DR   EMBL; AJ409219; CAC34727.1; -.
DR   HSSP; Q88571; 1TME.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     213    213
SQ   SEQUENCE   213 AA;  23494 MW;  92CBE4DBC885F3C9 CRC64;

  Query Match             67.3%;  Score 68;  DB 12;  Length 213;
  Best Local Similarity   78.9%;  Pred. No. 0.0052;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              ||  ||| |||| ||||||
Db        140 GSARRGDMGSLAARVARQL 158


RESULT 7
Q67438
ID   Q67438      PRELIMINARY;      PRT;   210 AA.
AC   Q67438; Q67439;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Foot and mouth disease virus (Serotype A24), capsid protein VP1 mRNA
DE   (Fragment).
OS   Foot-and-mouth disease virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12110;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85190547; PubMed=2986125;
RA   Weddell G.N., Yansura D.G., Dowbenko D.J., Hoatlin M.E., Grubman M.J.,
RA   Moore D.M., Kleid D.G.;
RT   "Sequence variation in the gene for the immunogenic capsid protein VP1
RT   of foot-and-mouth disease virus type A.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:2618-2622(1985).
DR   EMBL; K03340; AAA42595.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     210    210
SQ   SEQUENCE   210 AA;  23089 MW;  1C96CA1C5E4F6A60 CRC64;

  Query Match             66.3%;  Score 67;  DB 12;  Length 210;
  Best Local Similarity   73.7%;  Pred. No. 0.0075;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              ||| ||| |||| || :||
Db        137 GSGRRGDMGSLAARVVKQL 155


RESULT 8
Q9Q2N6
ID   Q9Q2N6      PRELIMINARY;      PRT;   969 AA.
AC   Q9Q2N6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Polyprotein (Fragment).
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A24/Cruzeiro/Brazil/55;
RA   Abrams C.C.;
RT   "Nucleotide sequence of foot-and-mouth disease virus
RT   A24/Cruzeiro/Brazil/55 from the poly(C) tract to 2B.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A24/Cruzeiro/Brazil/55;
RA   Knowles N.J., Samuel A.R., Aktas S., Rowe C.A., Abrams C.C.,
RA   Newman J.W.I., King A.M.Q.;
RT   "Phylogenetic comparison of the capsid-coding region of all seven
RT   foot-and-mouth disease virus serotypes.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ251476; CAB62583.1; -.
DR   HSSP; Q88571; 1TME.
DR   MEROPS; C28.001; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 3.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   CHAIN         1   >201       LEADER.
FT   CHAIN       202   >286       1A (VP4).
FT   CHAIN       287   >504       1B (VP2).
FT   CHAIN       505   >725       1C (VP3).
FT   CHAIN       726   >938       1D (VP1).
FT   CHAIN       939   >954       2A.
FT   CHAIN       955   >969       2B.
FT   NON_TER     969    969
SQ   SEQUENCE   969 AA;  107629 MW;  7DD5D908FCFF89FD CRC64;

  Query Match             66.3%;  Score 67;  DB 12;  Length 969;
  Best Local Similarity   73.7%;  Pred. No. 0.041;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              ||| ||| |||| || :||
Db        865 GSGRRGDMGSLAARVVKQL 883


RESULT 9
Q67442
ID   Q67442      PRELIMINARY;      PRT;   169 AA.
AC   Q67442; Q67443;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Foot and mouth disease virus (Serotype A32), capsid protein VP1 mRNA
DE   (Fragment).
OS   Foot-and-mouth disease virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12110;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85190547; PubMed=2986125;
RA   Weddell G.N., Yansura D.G., Dowbenko D.J., Hoatlin M.E., Grubman M.J.,
RA   Moore D.M., Kleid D.G.;
RT   "Sequence variation in the gene for the immunogenic capsid protein VP1
RT   of foot-and-mouth disease virus type A.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:2618-2622(1985).
DR   EMBL; K03342; AAA42598.1; -.
DR   HSSP; Q88571; 1TME.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 2.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     169    169
SQ   SEQUENCE   169 AA;  18499 MW;  9E8AEDC94AC57F67 CRC64;

  Query Match             65.3%;  Score 66;  DB 12;  Length 169;
  Best Local Similarity   73.7%;  Pred. No. 0.0085;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              ||| ||| | || |||:||
Db         96 GSGRRGDLGPLAARVAKQL 114


RESULT 10
Q8V443
ID   Q8V443      PRELIMINARY;      PRT;   126 AA.
AC   Q8V443;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   VP1 (Fragment).
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KEN/1/76;
RA   van Rensburg H.G., Bastos A.D., Haydon D., Joubert F., Nel L.H.;
RT   "Gene heterogeneity in the leader and 3C proteinases of foot-and-mouth
RT   disease virus.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF390861; AAL73359.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     126    126
SQ   SEQUENCE   126 AA;  13693 MW;  82675FAC45F713BB CRC64;

  Query Match             64.4%;  Score 65;  DB 12;  Length 126;
  Best Local Similarity   77.8%;  Pred. No. 0.009;
  Matches   14;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 SGVRGDFGSLALRVARQL 20
              || ||| |:|| ||||||
Db         54 SGRRGDMGALAARVARQL 71


RESULT 11
Q8V442
ID   Q8V442      PRELIMINARY;      PRT;   125 AA.
AC   Q8V442;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   VP1 (Fragment).
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GAM/51/98;
RA   van Rensburg H.G., Bastos A.D., Haydon D., Joubert F., Nel L.H.;
RT   "Gene heterogeneity in the leader and 3C proteinases of foot-and-mouth
RT   disease virus.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF390862; AAL73360.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     125    125
SQ   SEQUENCE   125 AA;  13486 MW;  2D172B9F0C281A02 CRC64;

  Query Match             61.4%;  Score 62;  DB 12;  Length 125;
  Best Local Similarity   73.7%;  Pred. No. 0.027;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              ||  ||| ||||:||| ||
Db         53 GSTRRGDLGSLAVRVAAQL 71


RESULT 12
Q9YPU0
ID   Q9YPU0      PRELIMINARY;      PRT;   143 AA.
AC   Q9YPU0;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   VP1 protein (Fragment).
GN   VP1.
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NM/XZ/64;
RA   Liu Z., Zhao Q., Liu W., Xie Q.;
RT   "Analysis of VP1 coding nucleotide sequences of six strains of foot-
RT   and-mouth disease virus type A.";
RL   Ping Tu Hsueh Pao 14:60-67(1998).
DR   EMBL; AJ131664; CAA10474.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     143    143
SQ   SEQUENCE   143 AA;  15881 MW;  CF0CC1858CC331E0 CRC64;

  Query Match             61.4%;  Score 62;  DB 12;  Length 143;
  Best Local Similarity   72.2%;  Pred. No. 0.032;
  Matches   13;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 SGVRGDFGSLALRVARQL 20
              :| ||| |||| |||:||
Db         71 AGRRGDLGSLAARVAKQL 88


RESULT 13
Q8JUR8
ID   Q8JUR8      PRELIMINARY;      PRT;   208 AA.
AC   Q8JUR8;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
```

```
DE   1D protein (Fragment).
GN   POL.
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/IND/170/88;
RX   MEDLINE=21954626; PubMed=11958451;
RA   Tosh C., Sanyal A., Hemadri D., Venkataramanan R.;
RT   "Phylogenetic analysis of serotype A foot-and-mouth disease virus
RT   isolated in India between 1977 and 2000.";
RL   Arch. Virol. 147:493-513(2002).
DR   EMBL; AF390612; AAM63984.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  22607 MW;  6CE616F86E60CF7E CRC64;

  Query Match             61.4%;  Score 62;  DB 12;  Length 208;
  Best Local Similarity   68.4%;  Pred. No. 0.048;
  Matches   13;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              |:| ||| |:|| ||| ||
Db        140 GAGRRGDLGALAARVAAQL 158


RESULT 14
Q8JUP0
ID   Q8JUP0      PRELIMINARY;      PRT;   211 AA.
AC   Q8JUP0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   1D protein (Fragment).
GN   POL.
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/IND/302/88;
RX   MEDLINE=21954626; PubMed=11958451;
RA   Tosh C., Sanyal A., Hemadri D., Venkataramanan R.;
RT   "Phylogenetic analysis of serotype A foot-and-mouth disease virus
RT   isolated in India between 1977 and 2000.";
RL   Arch. Virol. 147:493-513(2002).
DR   EMBL; AF390641; AAM64012.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     211    211
SQ   SEQUENCE   211 AA;  23100 MW;  CFBFC94CCD9E2FE3 CRC64;

  Query Match             61.4%;  Score 62;  DB 12;  Length 211;
  Best Local Similarity   68.4%;  Pred. No. 0.049;
  Matches   13;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              |:| ||| |:|| ||| ||
Db        140 GTGRRGDLGALAARVAAQL 158


RESULT 15
Q8JUP1
ID   Q8JUP1      PRELIMINARY;      PRT;   212 AA.
AC   Q8JUP1;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   1D protein (Fragment).
GN   POL.
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/IND/299/99;
RX   MEDLINE=21954626; PubMed=11958451;
RA   Tosh C., Sanyal A., Hemadri D., Venkataramanan R.;
RT   "Phylogenetic analysis of serotype A foot-and-mouth disease virus
RT   isolated in India between 1977 and 2000.";
RL   Arch. Virol. 147:493-513(2002).
DR   EMBL; AF390640; AAM64011.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     212    212
SQ   SEQUENCE   212 AA;  23150 MW;  E6B690890042CD3A CRC64;

  Query Match             61.4%;  Score 62;  DB 12;  Length 212;
  Best Local Similarity   68.4%;  Pred. No. 0.049;
  Matches   13;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              |:| ||| |:|| ||| ||
Db        139 GTGRRGDLGALAARVAAQL 157


Search completed: October  9, 2003, 10:17:34
Job time : 49.3333 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:11:24 ; Search time 17.3333 Seconds
(without alignments)
48.820 Million cell updates/sec

Title: US-09-549-186B-8
Perfect score: 101
Sequence: 1 CGSGVRGDFGSLALRVARQL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	101	100.0	20	4	US-08-716-249-8	Sequence 8, Appli
2	94	93.1	20	4	US-08-716-249-7	Sequence 7, Appli
3	93	92.1	20	4	US-08-716-249-9	Sequence 9, Appli
4	85	84.2	35	3	US-09-100-600A-1	Sequence 1, Appli
5	85	84.2	46	3	US-09-100-600A-38	Sequence 38, Appl
6	85	84.2	75	3	US-09-100-600A-28	Sequence 28, Appl
7	83	82.2	31	1	US-08-418-716A-3	Sequence 3, Appli
8	82	81.2	31	1	US-08-418-716A-4	Sequence 4, Appli
9	81	80.2	31	1	US-08-418-716A-5	Sequence 5, Appli
10	81	80.2	31	1	US-08-418-716A-6	Sequence 6, Appli
11	79.5	78.7	19	4	US-08-716-249-2	Sequence 2, Appli
12	79	78.2	31	1	US-08-418-716A-2	Sequence 2, Appli
13	78.5	77.7	19	4	US-08-716-249-1	Sequence 1, Appli
14	78	77.2	31	1	US-08-418-716A-1	Sequence 1, Appli
15	78	77.2	31	1	US-08-418-716A-7	Sequence 7, Appli
16	77	76.2	35	3	US-09-100-600A-2	Sequence 2, Appli
17	77	76.2	46	3	US-09-100-600A-3	Sequence 3, Appli
18	77	76.2	65	3	US-09-100-600A-27	Sequence 27, Appl
19	77	76.2	75	3	US-09-100-600A-29	Sequence 29, Appl
20	71	70.3	35	3	US-09-100-600A-8	Sequence 8, Appli
21	71	70.3	46	3	US-09-100-600A-42	Sequence 42, Appl
22	71	70.3	46	3	US-09-100-600A-49	Sequence 49, Appl
23	71	70.3	46	3	US-09-100-600A-50	Sequence 50, Appl
24	67	66.3	46	3	US-09-100-600A-43	Sequence 43, Appl
25	67	66.3	46	3	US-09-100-600A-44	Sequence 44, Appl
26	67	66.3	46	3	US-09-100-600A-48	Sequence 48, Appl
27	66	65.3	46	3	US-09-100-600A-45	Sequence 45, Appl
28	63	62.4	35	3	US-09-100-600A-9	Sequence 9, Appli
29	63	62.4	55	3	US-09-100-600A-33	Sequence 33, Appl
30	60	59.4	46	3	US-09-100-600A-39	Sequence 39, Appl
31	59	58.4	46	3	US-09-100-600A-46	Sequence 46, Appl
32	55	54.5	45	3	US-09-100-600A-55	Sequence 55, Appl
33	54	53.5	46	3	US-09-100-600A-47	Sequence 47, Appl
34	53	52.5	45	3	US-09-100-600A-60	Sequence 60, Appl
35	50	49.5	45	3	US-09-100-600A-51	Sequence 51, Appl
36	50	49.5	45	3	US-09-100-600A-62	Sequence 62, Appl
37	50	49.5	45	3	US-09-100-600A-64	Sequence 64, Appl
38	50	49.5	45	3	US-09-100-600A-65	Sequence 65, Appl
39	50	49.5	46	3	US-09-100-600A-40	Sequence 40, Appl
40	50	49.5	46	3	US-09-100-600A-41	Sequence 41, Appl
41	48	47.5	45	3	US-09-100-600A-56	Sequence 56, Appl
42	47	46.5	170	4	US-09-732-210-1676	Sequence 1676, Ap
43	47	46.5	387	4	US-09-252-991A-17881	Sequence 17881, A
44	46	45.5	45	3	US-09-100-600A-52	Sequence 52, Appl
45	46	45.5	54	3	US-09-100-600A-34	Sequence 34, Appl

ALIGNMENTS

```
RESULT 1
US-08-716-249-8
; Sequence 8, Application US/08716249
; Patent No. 6455244
;  GENERAL INFORMATION:
;    APPLICANT:  Guichard, Gilles
;    APPLICANT:  Muller, Sylviane
;    APPLICANT:  Briand, Jean-Paul
;    APPLICANT:  Regenmortel, Marc
;    TITLE OF INVENTION:  Retropeptides, Antibodies Thereto, and
;    TITLE OF INVENTION:  Uses Thereof for Vaccination and In Vitro Diagnosis
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spencer & Frank
;      STREET:  1100 New York Avenue, Suite 300E
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/716,249
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/00292
;      FILING DATE:  13-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Calvetti, Frederick F.
;      REGISTRATION NUMBER:  28,557
;      REFERENCE/DOCKET NUMBER:  GROFO 7001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)414-4000
;      TELEFAX:  (202)414-4040
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  YES
US-08-716-249-8

  Query Match           100.0%;  Score 101;  DB 4;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 2.1e-09;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 CGSGVRGDFGSLALRVARQL 20
              ||||||||||||||||||||
Db          1 CGSGVRGDFGSLALRVARQL 20


RESULT 2
US-08-716-249-7
; Sequence 7, Application US/08716249
; Patent No. 6455244
;  GENERAL INFORMATION:
;    APPLICANT:  Guichard, Gilles
;    APPLICANT:  Muller, Sylviane
;    APPLICANT:  Briand, Jean-Paul
;    APPLICANT:  Regenmortel, Marc
;    TITLE OF INVENTION:  Retropeptides, Antibodies Thereto, and
;    TITLE OF INVENTION:  Uses Thereof for Vaccination and In Vitro Diagnosis
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spencer & Frank
;      STREET:  1100 New York Avenue, Suite 300E
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/716,249
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/00292
;      FILING DATE:  13-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Calvetti, Frederick F.
;      REGISTRATION NUMBER:  28,557
;      REFERENCE/DOCKET NUMBER:  GROFO 7001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)414-4000
;      TELEFAX:  (202)414-4040
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-716-249-7

  Query Match             93.1%;  Score 94;  DB 4;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 2.6e-08;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDFGSLALRVARQL 20
              |||||||||||||| ||||||
Db          1 CGSGVRGDFGSLAPRVARQL 20


RESULT 3
US-08-716-249-9
; Sequence 9, Application US/08716249
; Patent No. 6455244
;  GENERAL INFORMATION:
;    APPLICANT:  Guichard, Gilles
;    APPLICANT:  Muller, Sylviane
;    APPLICANT:  Briand, Jean-Paul
;    APPLICANT:  Regenmortel, Marc
;    TITLE OF INVENTION:  Retropeptides, Antibodies Thereto, and
;    TITLE OF INVENTION:  Uses Thereof for Vaccination and In Vitro Diagnosis
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spencer & Frank
;      STREET:  1100 New York Avenue, Suite 300E
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/716,249
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/00292
;      FILING DATE:  13-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Calvetti, Frederick F.
;      REGISTRATION NUMBER:  28,557
;      REFERENCE/DOCKET NUMBER:  GROFO 7001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)414-4000
;      TELEFAX:  (202)414-4040
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  YES
US-08-716-249-9

  Query Match             92.1%;  Score 93;  DB 4;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 3.7e-08;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDFGSLALRVARQL 20
              |||||||| ||||||||||
Db          1 CGSGVRGDSGSLALRVARQL 20


RESULT 4
US-09-100-600A-1
; Sequence 1, Application US/09100600A
; Patent No. 6107021
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Shen, Ming
;    TITLE OF INVENTION:  SYNTHETIC PEPTIDE VACCINES FOR
;    TITLE OF INVENTION:  FOOT-AND-MOUTH DISEASE
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,600A
;      FILING DATE:  20-Jun-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria C.H. Lin
```

```
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4156
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  35 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-600A-1

  Query Match             84.2%;  Score 85;  DB 3;  Length 35;
  Best Local Similarity   94.7%;  Pred. No. 1.2e-06;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              |||||||||||| ||||||
Db          7 GSGVRGDFGSLAPRVARQL 25


RESULT 5
US-09-100-600A-38
; Sequence 38, Application US/09100600A
; Patent No. 6107021
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Shen, Ming
;    TITLE OF INVENTION:  SYNTHETIC PEPTIDE VACCINES FOR
;    TITLE OF INVENTION:  FOOT-AND-MOUTH DISEASE
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,600A
;      FILING DATE:  20-Jun-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria C.H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4156
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-600A-38

  Query Match             84.2%;  Score 85;  DB 3;  Length 46;
  Best Local Similarity   94.7%;  Pred. No. 1.6e-06;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              |||||||||||| ||||||
Db         16 GSGVRGDFGSLAPRVARQL 34


RESULT 6
```

```
US-09-100-600A-28
; Sequence 28, Application US/09100600A
; Patent No. 6107021
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Shen, Ming
;    TITLE OF INVENTION:  SYNTHETIC PEPTIDE VACCINES FOR
;    TITLE OF INVENTION:  FOOT-AND-MOUTH DISEASE
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,600A
;      FILING DATE:  20-Jun-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria C.H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4156
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  75 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-600A-28

  Query Match             84.2%;  Score 85;  DB 3;  Length 75;
  Best Local Similarity   94.7%;  Pred. No. 2.6e-06;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              |||||||||||| ||||||
Db         47 GSGVRGDFGSLAPRVARQL 65


RESULT 7
US-08-418-716A-3
; Sequence 3, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;    APPLICANT:  Mason, Peter W.
;    APPLICANT:  Baxt, Barry
;    APPLICANT:  Reider, Elizabeth
;    APPLICANT:  Berinstein, Analia
;    APPLICANT:  Kang, Angray S
;    TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;    TITLE OF INVENTION:  Viruses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janelle S. Graeter
;      STREET:  Rm. 411, Bldg. 005, BARC-W
;      CITY:  Beltsville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20705
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/418,716A
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Graeter, Janelle S.
;       REGISTRATION NUMBER:  35,024
;       REFERENCE/DOCKET NUMBER:  0137.94
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 504-6629
;       TELEFAX:  (301) 504-5060
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  31 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     HYPOTHETICAL:  NO
;     FRAGMENT TYPE:  internal
;     ORIGINAL SOURCE:
;       ORGANISM:  Foot and mouth disease virus
US-08-418-716A-3

  Query Match             82.2%;  Score 83;  DB 1;  Length 31;
  Best Local Similarity   94.7%;  Pred. No. 2.1e-06;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              ||||||| ||||||||||
Db         11 GSGVRGDKGSLALRVARQL 29


RESULT 8
US-08-418-716A-4
; Sequence 4, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;    APPLICANT:  Mason, Peter W.
;    APPLICANT:  Baxt, Barry
;    APPLICANT:  Reider, Elizabeth
;    APPLICANT:  Berinstein, Analia
;    APPLICANT:  Kang, Angray S
;    TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;    TITLE OF INVENTION:  Viruses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janelle S. Graeter
;      STREET:  Rm. 411, Bldg. 005, BARC-W
;      CITY:  Beltsville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20705
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/418,716A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Graeter, Janelle S.
;      REGISTRATION NUMBER:  35,024
;      REFERENCE/DOCKET NUMBER:  0137.94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 504-6629
;      TELEFAX:  (301) 504-5060
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:
;      ORGANISM:  Foot and mouth disease virus
US-08-418-716A-4

  Query Match             81.2%;  Score 82;  DB 1;  Length 31;
  Best Local Similarity   89.5%;  Pred. No. 3e-06;
  Matches   17;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              ||||:||||||| ||||||
Db         11 GSGVKGDFGSLAPRVARQL 29


RESULT 9
US-08-418-716A-5
; Sequence 5, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;    APPLICANT:  Mason, Peter W.
;    APPLICANT:  Baxt, Barry
;    APPLICANT:  Reider, Elizabeth
;    APPLICANT:  Berinstein, Analia
;    APPLICANT:  Kang, Angray S
;    TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;    TITLE OF INVENTION:  Viruses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janelle S. Graeter
;      STREET:  Rm. 411, Bldg. 005, BARC-W
;      CITY:  Beltsville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20705
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/418,716A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Graeter, Janelle S.
;      REGISTRATION NUMBER:  35,024
;      REFERENCE/DOCKET NUMBER:  0137.94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 504-6629
;      TELEFAX:  (301) 504-5060
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:
;      ORGANISM:  Foot and mouth disease virus
US-08-418-716A-5

  Query Match             80.2%;  Score 81;  DB 1;  Length 31;
  Best Local Similarity   89.5%;  Pred. No. 4.3e-06;
  Matches   17;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
```

```
                  ||||||:||||| ||||||
Db         11 GSGVRGEFGSLAPRVARQL 29


RESULT 10
US-08-418-716A-6
; Sequence 6, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;    APPLICANT:  Mason, Peter W.
;    APPLICANT:  Baxt, Barry
;    APPLICANT:  Reider, Elizabeth
;    APPLICANT:  Berinstein, Analia
;    APPLICANT:  Kang, Angray S
;    TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;    TITLE OF INVENTION:  Viruses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janelle S. Graeter
;      STREET:  Rm. 411, Bldg. 005, BARC-W
;      CITY:  Beltsville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20705
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/418,716A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Graeter, Janelle S.
;      REGISTRATION NUMBER:  35,024
;      REFERENCE/DOCKET NUMBER:  0137.94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 504-6629
;      TELEFAX:  (301) 504-5060
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:
;      ORGANISM:  Foot and mouth disease virus
US-08-418-716A-6

  Query Match             80.2%;  Score 81;  DB 1;  Length 31;
  Best Local Similarity   89.5%;  Pred. No. 4.3e-06;
  Matches   17;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              ||||||:||||| ||||||
Db         11 GSGVRGEFGSLAPRVARQL 29


RESULT 11
US-08-716-249-2
; Sequence 2, Application US/08716249
; Patent No. 6455244
;  GENERAL INFORMATION:
;    APPLICANT:  Guichard, Gilles
;    APPLICANT:  Muller, Sylviane
;    APPLICANT:  Briand, Jean-Paul
;    APPLICANT:  Regenmortel, Marc
;    TITLE OF INVENTION:  Retropeptides, Antibodies Thereto, and
;    TITLE OF INVENTION:  Uses Thereof for Vaccination and In Vitro Diagnosis
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spencer & Frank
;      STREET:  1100 New York Avenue, Suite 300E
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/716,249
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/00292
;      FILING DATE:  13-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Calvetti, Frederick F.
;      REGISTRATION NUMBER:  28,557
;      REFERENCE/DOCKET NUMBER:  GROFO 7001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)414-4000
;      TELEFAX:  (202)414-4040
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-716-249-2

  Query Match             78.7%;  Score 79.5;  DB 4;  Length 19;
  Best Local Similarity   90.0%;  Pred. No. 4.4e-06;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 CGSGVRGDFGSLALRVARQL 20
              |||||||||||| | ||||||
Db          1 CGSGVRGDFGS-APRVARQL 19


RESULT 12
US-08-418-716A-2
; Sequence 2, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;    APPLICANT:  Mason, Peter W.
;    APPLICANT:  Baxt, Barry
;    APPLICANT:  Reider, Elizabeth
;    APPLICANT:  Berinstein, Analia
;    APPLICANT:  Kang, Angray S
;    TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;    TITLE OF INVENTION:  Viruses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janelle S. Graeter
;      STREET:  Rm. 411, Bldg. 005, BARC-W
;      CITY:  Beltsville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20705
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/418,716A
;      FILING DATE:
```

```
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Graeter, Janelle S.
;      REGISTRATION NUMBER:  35,024
;      REFERENCE/DOCKET NUMBER:  0137.94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 504-6629
;      TELEFAX:  (301) 504-5060
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:
;      ORGANISM:  Foot and mouth disease virus
US-08-418-716A-2

  Query Match             78.2%;  Score 79;  DB 1;  Length 31;
  Best Local Similarity   89.5%;  Pred. No. 8.8e-06;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              ||| |||||||| ||||||
Db         11 GSGPRGDFGSLAPRVARQL 29


RESULT 13
US-08-716-249-1
; Sequence 1, Application US/08716249
; Patent No. 6455244
;  GENERAL INFORMATION:
;    APPLICANT:  Guichard, Gilles
;    APPLICANT:  Muller, Sylviane
;    APPLICANT:  Briand, Jean-Paul
;    APPLICANT:  Regenmortel, Marc
;    TITLE OF INVENTION:  Retropeptides, Antibodies Thereto, and
;    TITLE OF INVENTION:  Uses Thereof for Vaccination and In Vitro Diagnosis
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spencer & Frank
;      STREET:  1100 New York Avenue, Suite 300E
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/716,249
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/00292
;      FILING DATE:  13-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Calvetti, Frederick F.
;      REGISTRATION NUMBER:  28,557
;      REFERENCE/DOCKET NUMBER:  GROFO 7001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)414-4000
;      TELEFAX:  (202)414-4040
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-716-249-1

  Query Match             77.7%;  Score 78.5;  DB 4;  Length 19;
  Best Local Similarity   90.0%;  Pred. No. 6.3e-06;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 CGSGVRGDFGSLALRVARQL 20
              |||||||| || ||||||||
Db          1 CGSGVRGDSGS-ALRVARQL 19


RESULT 14
US-08-418-716A-1
; Sequence 1, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;    APPLICANT:  Mason, Peter W.
;    APPLICANT:  Baxt, Barry
;    APPLICANT:  Reider, Elizabeth
;    APPLICANT:  Berinstein, Analia
;    APPLICANT:  Kang, Angray S
;    TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;    TITLE OF INVENTION:  Viruses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janelle S. Graeter
;      STREET:  Rm. 411, Bldg. 005, BARC-W
;      CITY:  Beltsville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20705
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/418,716A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Graeter, Janelle S.
;      REGISTRATION NUMBER:  35,024
;      REFERENCE/DOCKET NUMBER:  0137.94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 504-6629
;      TELEFAX:  (301) 504-5060
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:
;      ORGANISM:  Foot and mouth disease virus
US-08-418-716A-1

  Query Match             77.2%;  Score 78;  DB 1;  Length 31;
  Best Local Similarity   89.5%;  Pred. No. 1.3e-05;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              ||| |||||||| ||||||
Db         11 GSGDRGDFGSLAPRVARQL 29


RESULT 15
US-08-418-716A-7
```

```
; Sequence 7, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;    APPLICANT:  Mason, Peter W.
;    APPLICANT:  Baxt, Barry
;    APPLICANT:  Reider, Elizabeth
;    APPLICANT:  Berinstein, Analia
;    APPLICANT:  Kang, Angray S
;    TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;    TITLE OF INVENTION:  Viruses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janelle S. Graeter
;      STREET:  Rm. 411, Bldg. 005, BARC-W
;      CITY:  Beltsville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20705
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/418,716A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Graeter, Janelle S.
;      REGISTRATION NUMBER:  35,024
;      REFERENCE/DOCKET NUMBER:  0137.94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 504-6629
;      TELEFAX:  (301) 504-5060
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:
;      ORGANISM:  Foot and mouth disease virus
US-08-418-716A-7

  Query Match             77.2%;  Score 78;  DB 1;  Length 31;
  Best Local Similarity   84.2%;  Pred. No. 1.3e-05;
  Matches   16;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              ||||:|:||||| ||||||
Db         11 GSGVKGEFGSLAPRVARQL 29


Search completed: October  9, 2003, 10:19:38
Job time : 18.3333 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:14:31 ; Search time 21.3333 Seconds
(without alignments)
151.058 Million cell updates/sec

Title: US-09-549-186B-8
Perfect score: 101
Sequence: 1 CGSGVRGDFGSLALRVARQL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	47	46.5	173	15	US-10-156-761-12492	Sequence 12492, A
2	43	42.6	261	15	US-10-156-761-9294	Sequence 9294, Ap
3	43	42.6	294	15	US-10-102-806-595	Sequence 595, App
4	43	42.6	312	15	US-10-207-655-186	Sequence 186, App
5	42	41.6	350	11	US-09-847-208-53	Sequence 53, Appl
6	42	41.6	476	9	US-09-291-299A-3	Sequence 3, Appli
7	42	41.6	509	14	US-10-072-152-6	Sequence 6, Appli
8	42	41.6	627	15	US-10-222-100-3	Sequence 3, Appli
9	41.5	41.1	446	10	US-09-738-626-6262	Sequence 6262, Ap
10	41	40.6	74	15	US-10-106-698-7642	Sequence 7642, Ap
11	41	40.6	337	9	US-09-815-242-11926	Sequence 11926, A
12	41	40.6	394	15	US-10-156-761-14684	Sequence 14684, A
13	41	40.6	648	12	US-10-128-713A-2	Sequence 2, Appli
14	41	40.6	722	14	US-10-128-870-23	Sequence 23, Appl
15	41	40.6	722	15	US-10-131-685-23	Sequence 23, Appl
16	40.5	40.1	378	12	US-10-166-225A-49	Sequence 49, Appl
17	40	39.6	94	15	US-10-102-806-581	Sequence 581, App
18	40	39.6	142	9	US-09-908-711-105	Sequence 105, App
19	40	39.6	142	11	US-09-764-891-3245	Sequence 3245, Ap
20	40	39.6	283	15	US-10-156-761-9949	Sequence 9949, Ap
21	40	39.6	559	9	US-09-988-200-2	Sequence 2, Appli
22	40	39.6	1150	10	US-09-870-122-3	Sequence 3, Appli
23	40	39.6	1164	10	US-09-870-122-1	Sequence 1, Appli
24	40	39.6	1167	10	US-09-870-122-2	Sequence 2, Appli
25	40	39.6	1181	10	US-09-870-122-23	Sequence 23, Appl
26	40	39.6	4123	15	US-10-213-509-5	Sequence 5, Appli
27	39	38.6	28	9	US-09-864-761-34511	Sequence 34511, A
28	39	38.6	440	14	US-10-016-283-26	Sequence 26, Appl
29	39	38.6	456	14	US-10-016-283-25	Sequence 25, Appl
30	39	38.6	492	14	US-10-016-283-36	Sequence 36, Appl
31	39	38.6	508	11	US-09-895-298-92	Sequence 92, Appl
32	39	38.6	541	9	US-09-815-242-5807	Sequence 5807, Ap
33	39	38.6	602	9	US-09-815-242-12896	Sequence 12896, A
34	39	38.6	1014	9	US-09-912-020-266	Sequence 266, App
35	39	38.6	1105	15	US-10-128-714-3447	Sequence 3447, Ap
36	39	38.6	1208	15	US-10-156-761-13251	Sequence 13251, A
37	38.5	38.1	47	9	US-09-864-761-47530	Sequence 47530, A
38	38.5	38.1	819	11	US-09-866-050A-503	Sequence 503, App
39	38	37.6	19	11	US-09-931-325A-145	Sequence 145, App
40	38	37.6	19	12	US-09-930-915A-186	Sequence 186, App
41	38	37.6	26	11	US-09-931-325A-146	Sequence 146, App
42	38	37.6	26	12	US-09-930-915A-187	Sequence 187, App
43	38	37.6	31	12	US-09-930-915A-7	Sequence 7, Appli
44	38	37.6	31	12	US-10-082-014-13	Sequence 13, Appl
45	38	37.6	99	15	US-10-106-698-4862	Sequence 4862, Ap

ALIGNMENTS

RESULT 1
US-10-156-761-12492
; Sequence 12492, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12492
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12492

Query Match 46.5%; Score 47; DB 15; Length 173;
Best Local Similarity 57.9%; Pred. No. 7.9;
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```
Qy          2 GSGVRGDFGSLALRVARQL 20
              | || |  |:| | ||| |
Db        111 GGGVSGQAGALRLGVARAL 129
```

RESULT 2
US-10-156-761-9294

; Sequence 9294, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9294
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9294

Query Match 42.6%; Score 43; DB 15; Length 261;
Best Local Similarity 52.6%; Pred. No. 54;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 2 GSGVRGDFGSLALRVARQL 20
Db 97 GQSVLLDIGTTAYRLARQL 115

RESULT 3
US-10-102-806-595
; Sequence 595, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 595
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (269)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (278)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-595

Query Match 42.6%; Score 43; DB 15; Length 294;
Best Local Similarity 53.3%; Pred. No. 61;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 3 SGVRGDFGSLALRVA 17
Db 159 SGTSGDHGELVVRIA 173

RESULT 4
US-10-207-655-186
; Sequence 186, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 186
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-186

Query Match 42.6%; Score 43; DB 15; Length 312;
Best Local Similarity 61.5%; Pred. No. 65;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGSGVRGDFGSLA 13
Db 164 CGSGARGGDGEVA 176

RESULT 5
US-09-847-208-53
; Sequence 53, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; TITLE OF INVENTION: IgE-MEDIATED ALLERGIC DISEASES
; FILE REFERENCE: UC67.002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Candida albicans (Yeast)
US-09-847-208-53

Query Match 41.6%; Score 42; DB 11; Length 350;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 3 SGVRGDFGSLALRVARQL 20
Db 179 SGAGGGLGSLAVQYARAM 196

RESULT 6
US-09-291-299A-3
; Sequence 3, Application US/09291299A
; Patent No. US20020076405A1
; GENERAL INFORMATION:
; APPLICANT: Yong, Xie
; TITLE OF INVENTION: Antigenized Antibody Vaccine for Foot-and-Mouth Disease
; FILE REFERENCE: 09/291,299
; CURRENT APPLICATION NUMBER: US/09/291,299A
; CURRENT FILING DATE: 1999-04-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 476

```
;   TYPE: PRT
;   ORGANISM: Sus sp.
US-09-291-299A-3

  Query Match             41.6%;  Score 42;  DB 9;  Length 476;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+02;
  Matches   11;  Conservative    2;  Mismatches    7;  Indels    2;  Gaps    1;

Qy          1 CGS--GVRGDFGSLALRVARQL 20
              ||    :|||   || :||| |
Db        115 CGKVPNLRGDLQVLAQKVARTL 136


RESULT 7
US-10-072-152-6
; Sequence 6, Application US/10072152
; Publication No. US20020142438A1
; GENERAL INFORMATION:
;  APPLICANT: Andersen, Lene N.
;  APPLICANT:  Schulein, Martin
;  APPLICANT:  Lange, Niels E.
;  APPLICANT:  Bjornvad, Mads E.
;  APPLICANT:  Moller, Soren
;  APPLICANT:  Glad, Sanne O. S.
;  APPLICANT:  Kauppinen, Markus S.
;  APPLICANT:  Schnorr, Kirk
;  APPLICANT:  Kongsbak, Lars
;  TITLE OF INVENTION: No. US20020142438A1el Pectate Lyases
;  FILE REFERENCE: 5378.200-US
;  CURRENT APPLICATION NUMBER: US/10/072,152
;  CURRENT FILING DATE:  2002-02-07
;  PRIOR APPLICATION NUMBER: US/09/198,955
;  PRIOR FILING DATE: 1998-11-24
;  PRIOR APPLICATION NUMBER: 1343/97
;  PRIOR FILING DATE: 1997-11-24
;  PRIOR APPLICATION NUMBER: 1344/97
;  PRIOR FILING DATE: 1997-11-24
;  PRIOR APPLICATION NUMBER: 60/067,249
;  PRIOR FILING DATE: 1997-12-02
;  PRIOR APPLICATION NUMBER: 60/067,240
;  PRIOR FILING DATE: 1997-12-02
;  PRIOR APPLICATION NUMBER: 09/073,684
;  PRIOR FILING DATE: 1998-05-06
;  PRIOR APPLICATION NUMBER: 09/184,217
;  PRIOR FILING DATE: 1998-11-02
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 509
;   TYPE: PRT
;   ORGANISM: Bacillus sp.
US-10-072-152-6

  Query Match             41.6%;  Score 42;  DB 14;  Length 509;
  Best Local Similarity   46.7%;  Pred. No. 1.6e+02;
  Matches    7;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRV 16
              | | ||:|  : ||:
Db        266 GQGTRGEFDGIGLRL 280


RESULT 8
US-10-222-100-3
; Sequence 3, Application US/10222100
; Publication No. US20030059439A1
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  APPLICANT:  Revets, Hilde
;  APPLICANT:  Cornelis, Pierre
;  APPLICANT:  De Baetselier, Patrick
;  TITLE OF INVENTION: Th1 INDUCING NATURAL ADJUVANT FOR HETEROLOGOUS ANTIGENS
;  FILE REFERENCE: 2676-5449US
;  CURRENT APPLICATION NUMBER: US/10/222,100
;  CURRENT FILING DATE:  2002-08-16
;  PRIOR APPLICATION NUMBER: PCT/EP01/01673
;  PRIOR FILING DATE: 2000-02-18
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 627
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: OprI-3D-FMDV15 fusion
US-10-222-100-3

  Query Match             41.6%;  Score 42;  DB 15;  Length 627;
  Best Local Similarity   55.6%;  Pred. No. 2e+02;
  Matches   10;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 SGVRGDFGSLALRVARQL 20
              | :|||   || :||| |
Db        573 SNLRGDLQVLAQKVARTL 590


RESULT 9
US-09-738-626-6262
; Sequence 6262, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6262
;   LENGTH: 446
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-6262

  Query Match             41.1%;  Score 41.5;  DB 10;  Length 446;
  Best Local Similarity   55.0%;  Pred. No. 1.6e+02;
  Matches   11;  Conservative    2;  Mismatches    4;  Indels    3;  Gaps    1;

Qy          4 GVR---GDFGSLALRVARQL 20
              |||   || | || :| :||
Db        261 GVRIDSGDLGVLARKVRKQL 280


RESULT 10
US-10-106-698-7642
; Sequence 7642, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept
;  FILE REFERENCE: PA005P1
```

```
;  CURRENT APPLICATION NUMBER: US/10/106,698
;  CURRENT FILING DATE:  2002-03-27
;  PRIOR APPLICATION NUMBER: PCT/US00/26524
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: US 60/157,137
;  PRIOR FILING DATE: 1999-09-29
;  PRIOR APPLICATION NUMBER: US 60/163,280
;  PRIOR FILING DATE: 1999-11-03
;  NUMBER OF SEQ ID NOS: 8564
;  SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 7642
;   LENGTH: 74
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (49)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (54)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-7642

  Query Match            40.6%;  Score 41;  DB 15;  Length 74;
  Best Local Similarity   57.9%;  Pred. No. 29;
  Matches   11;  Conservative    1;  Mismatches    5;  Indels    2;  Gaps    1;

Qy          1 CGSG--VRGDFGSLALRVA 17
              |||   : |  ||||| ||
Db         26 CGSSRDLHGSQGSLALSVA 44


RESULT 11
US-09-815-242-11926
; Sequence 11926, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11926
;   LENGTH: 337
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11926

  Query Match            40.6%;  Score 41;  DB 9;  Length 337;
  Best Local Similarity   47.1%;  Pred. No. 1.5e+02;
  Matches    8;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 GVRGDFGSLALRVARQL 20
              |  |  ||: :::||||
Db        157 GAAGGVGSILVQLARQL 173


RESULT 12
US-10-156-761-14684
; Sequence 14684, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14684
;   LENGTH: 394
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-14684

  Query Match            40.6%;  Score 41;  DB 15;  Length 394;
  Best Local Similarity   52.6%;  Pred. No. 1.7e+02;
  Matches   10;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVARQL 20
              |||| ||  |  : :|| |
Db        192 GSGVLGDLASHGVDLARFL 210


RESULT 13
US-10-128-713A-2
; Sequence 2, Application US/10128713A
; Publication No. US20030170847A1
; GENERAL INFORMATION:
;  APPLICANT: Bramucci, Michael G
;  TITLE OF INVENTION: Genes Involved in Isoprenoid Compound Production
;  FILE REFERENCE: CL-1788
;  CURRENT APPLICATION NUMBER: US/10/128,713A
;  CURRENT FILING DATE:  2002-04-22
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
;   LENGTH: 648
;   TYPE: PRT
;   ORGANISM: Rhodococcus erythropolis
US-10-128-713A-2

  Query Match            40.6%;  Score 41;  DB 12;  Length 648;
  Best Local Similarity   44.0%;  Pred. No. 3e+02;
  Matches   11;  Conservative    2;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          4 GVRGD--------FGSLALRVARQL 20
              | |||        | |||| :| :|
Db        500 GERGDVLLVAVGPFASLALEIAERL 524


RESULT 14
US-10-128-870-23
```

```
; Sequence 23, Application US/10128870
; Publication No. US20020168724A1
; GENERAL INFORMATION:
;  APPLICANT: Blanar, Michael A.
;  APPLICANT:  Dworetzky, Steven
;  APPLICANT:  Gribkoff, Valentin K.
;  APPLICANT:  Levesque, Paul C.
;  APPLICANT:  Little, Wayne A.
;  APPLICANT:  Neubauer, Michael G.
;  APPLICANT:  Yang, Wen-Pin
;  TITLE OF INVENTION: KCNQ POTASSIUM CHANNELS AND METHODS OF MODULATING SAME
;  FILE REFERENCE: DC58aDIV
;  CURRENT APPLICATION NUMBER: US/10/128,870
;  CURRENT FILING DATE:  2002-04-24
;  PRIOR APPLICATION NUMBER: 09/105,058
;  PRIOR FILING DATE: June 26, 1998
;  PRIOR APPLICATION NUMBER: 60/055,599
;  PRIOR FILING DATE: August 12, 1997
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
;   LENGTH: 722
;   TYPE: PRT
;   ORGANISM: MOUSE
US-10-128-870-23

  Query Match             40.6%;  Score 41;  DB 14;  Length 722;
  Best Local Similarity   52.9%;  Pred. No. 3.3e+02;
  Matches    9;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVAR 18
              |:   || ||| ||: |
Db        698 GTSPVGDHGSLVLRLER 714


RESULT 15
US-10-131-685-23
; Sequence 23, Application US/10131685
; Publication No. US20030044912A1
; GENERAL INFORMATION:
;  APPLICANT: Blanar, Michael A.
;  APPLICANT:  Levesque, Paul C.
;  APPLICANT:  Little, Wayne A.
;  APPLICANT:  Neubauer, Michael G.
;  APPLICANT:  Yang, Wen-Pin
;  TITLE OF INVENTION: KCNQ POTASSIUM CHANNELS AND METHODS OF MODULATING SAME
;  FILE REFERENCE: DC58aCON
;  CURRENT APPLICATION NUMBER: US/10/131,685
;  CURRENT FILING DATE:  2002-07-23
;  PRIOR APPLICATION NUMBER: US 09/105,058
;  PRIOR FILING DATE: 1998-06-26
;  PRIOR APPLICATION NUMBER: US 60/055,599
;  PRIOR FILING DATE: 1997-08-12
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
;   LENGTH: 722
;   TYPE: PRT
;   ORGANISM: MOUSE
US-10-131-685-23

  Query Match             40.6%;  Score 41;  DB 15;  Length 722;
  Best Local Similarity   52.9%;  Pred. No. 3.3e+02;
  Matches    9;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDFGSLALRVAR 18
              |:   || ||| ||: |
Db        698 GTSPVGDHGSLVLRLER 714


Search completed: October  9, 2003, 10:20:48
Job time : 21.3333 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:02:18 ; Search time 58.3333 Seconds
(without alignments)
54.421 Million cell updates/sec

Title: US-09-549-186B-9
Perfect score: 99
Sequence: 1 CGSGVRGDSGSLALRVARQL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	90	90.9	20	14	AAR38543	Foot and Mouth Dis
2	90	90.9	20	14	AAR41793	Foot and Mouth Dis
3	86	86.9	31	18	AAW17085	Foot and mouth dis
4	84	84.8	19	24	ABU08643	Foot and mouth dis
5	77	77.8	19	24	ABU08642	Foot and mouth dis
6	77	77.8	20	21	AAY94587	Envelope gene epit
7	77	77.8	31	4	AAP30107	Sequence of VP1 ca
8	77	77.8	31	6	AAP50728	Foot-and-mouth dis
9	77	77.8	31	18	AAW17082	Foot and mouth dis
10	77	77.8	35	21	AAY68501	Target antigenic p
11	77	77.8	75	21	AAY68528	Synthetic foot and
12	77	77.8	216	4	AAP30205	Sequence encoded b
13	77	77.8	220	4	AAP30195	Sequence encoded b
14	77	77.8	233	4	AAP30204	Sequence encoded b
15	77	77.8	405	4	AAP30202	Sequence encoded b
16	77	77.8	406	4	AAP30203	Sequence encoded b
17	77	77.8	609	4	AAP30206	Sequence encoded b
18	74	74.7	31	18	AAW17086	Foot and mouth dis
19	73	73.7	31	18	AAW17087	Foot and mouth dis
20	71	71.7	31	18	AAW17084	Foot and mouth dis
21	70	70.7	31	18	AAW17083	Foot and mouth dis
22	70	70.7	31	18	AAW17088	Foot and mouth dis
23	70	70.7	213	4	AAP30198	Sequence encoded b
24	70	70.7	233	4	AAP30209	Sequence of antige
25	69	69.7	35	21	AAY68502	Target antigenic p
26	69	69.7	35	21	AAY68508	Consensus antigeni
27	69	69.7	46	21	AAY68503	Target antigenic p
28	69	69.7	65	21	AAY68527	Synthetic foot and
29	69	69.7	75	21	AAY68529	Synthetic foot and
30	67	67.7	35	7	AAP60895	Peptide containing
31	67	67.7	35	9	AAP82746	VP1(A12,119) contg
32	66	66.7	20	5	AAP40103	Sequence at antige
33	66	66.7	216	5	AAP40085	Sequence of foot a
34	63	63.6	248	4	AAP30207	Sequence of an FMD
35	62	62.6	18	4	AAP30110	Sequence of VP1 ca
36	61	61.6	35	21	AAY68509	Consensus antigeni
37	61	61.6	55	21	AAY68533	Synthetic foot and
38	57	57.6	218	4	AAP30196	Sequence encoded b
39	55	55.6	28	4	AAP30316	Sequence of VP1 ca
40	54	54.5	213	21	AAB15428	FMDV antigen for r
41	54	54.5	213	21	AAB15397	Foot-and-mouth dis
42	54	54.5	213	21	AAB15398	Foot-and-mouth dis
43	54	54.5	213	21	AAB15402	Foot-and-mouth dis
44	53	53.5	36	7	AAP60896	Peptide containing
45	53	53.5	36	9	AAP82747	VP1(A24,C) contg.

ALIGNMENTS

RESULT 1
AAR38543
ID AAR38543 standard; peptide; 20 AA.
XX
AC AAR38543;
XX
DT 25-MAR-2003 (updated)
DT 11-JAN-1994 (first entry)
XX
DE Foot and Mouth Disease Virus VPI antigenic peptide.
XX
KW Multivalent vaccine; polylysine; homopolymer; dendritic core;
KW multiple antigen peptide system; MAPS; Foot and Mouth Disease;
KW FMDV; VPI protein; Aphthovirus; vaccine.
XX
OS Foot and Mouth Disease Virus.
XX
PN US5229490-A.
XX
PD 20-JUL-1993.
XX
PF 20-DEC-1990; 90US-0631185.
XX
PR 06-MAY-1987; 87US-0047204.
PR 30-JUN-1987; 87US-0068840.
PR 12-APR-1989; 89US-0336845.
PR 20-DEC-1990; 90US-0631185.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI Tam JP;
XX

```
DR   WPI; 1993-242534/30.
XX
PT   Multiple antigenic peptide systems - have dendritic core attached
PT   covalently to antigens, used as vaccine
XX
PS   Claim 26; Column 24; 23pp; English.
XX
CC   This is a preferred FMDV antigenic peptide suitable for
CC   construction of a Multiple Antigen Peptide System (MAPS) by
CC   covalent coupling to a dendritic homopolymer core (esp.
CC   polylysine). The antigen is derived from the Foot and Mouth Disease
CC   Virus VPI protein.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence   20 AA;

  Query Match             90.9%;  Score 90;  DB 14;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-06;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              |||||||||||||||||||
Db          1 GSGVRGDSGSLALRVARQL 19


RESULT 2
AAR41793
ID   AAR41793 standard; protein; 20 AA.
XX
AC   AAR41793;
XX
DT   25-MAR-2003  (updated)
DT   25-MAR-1994  (first entry)
XX
DE   Foot and Mouth Disease Virus VPI antigen.
XX
KW   Group A Streptococci; surface antigen; anchor region;
KW   gram positive bacterium; M protein; fusion protein; immunogen;
KW   FMDV VPI; vaccine.
XX
OS   Foot and Mouth Disease Virus.
XX
PN   WO9318163-A2.
XX
PD   16-SEP-1993.
XX
PF   12-MAR-1993;   93WO-US02355.
XX
PR   13-MAR-1992;   92US-0851082.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Fischetti VA,  Pozzi G,  Schneewind O;
XX
DR   WPI; 1993-303477/38.
XX
PT   Gene encoding hybrid surface protein of gram positive bacteria -
PT   useful for preparing vaccine compsn. for protecting animals from
PT   bacterial infection
XX
PS   Disclosure; Page 55; 85pp; English.
XX
CC   Hybrid surface proteins are claimed in which an active
CC   polypeptide (e.g. a surface antigen from a mammalian tumour cell,
CC   sperm or an allergen, bacterium, virus, parasite or fungus) is
CC   fused to an anchor region from a surface antigen normally expressed
CC   on the cell surface of gram positive bacteria. The anchor segment
CC   is pref. derived from a streptococcal M protein (see AAR41780).
CC   Sequence AAR41793 is a peptide sequence from the VPI protein of
CC   Foot and Mouth Disease virus (Nature 298:30 (1983)) suitable for
CC   development of a vaccine against FMDV.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   20 AA;

  Query Match             90.9%;  Score 90;  DB 14;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-06;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              |||||||||||||||||||
Db          1 GSGVRGDSGSLALRVARQL 19


RESULT 3
AAW17085
ID   AAW17085 standard; peptide; 31 AA.
XX
AC   AAW17085;
XX
DT   13-JUN-1997  (first entry)
XX
DE   Foot and mouth disease virus modified G-H loop of protein VP-1.
XX
KW   Foot and mouth disease; aphthovirus; vaccine; antigen; epitope;
KW   immunogen; G-H loop; VP-1; deletion; mutant; binding site removal;
KW   antibody; non-infectious; attenuated.
XX
OS   Foot and mouth disease virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 18
FT                   /label= substitution
FT                   /note= "wild-type Phe replaced with Lys"
FT   Misc-difference 23
FT                   /label= substitution
FT                   /note= "wild-type Pro replaced with Leu"
XX
PN   US5612040-A.
XX
PD   18-MAR-1997.
XX
PF   07-APR-1995;   95US-0418716.
XX
PR   07-APR-1995;   95US-0418716.
XX
PA   (USDA ) US SEC OF AGRIC.
XX
PI   Baxt B,  Berinstein A,  Kang AS,  Mason PW,  Reider E;
XX
DR   WPI; 1997-192081/17.
XX
PT   Genetically modified foot-and-mouth disease virus - lacks
PT   cell-binding site and is non-infectious, useful in vaccines
XX
PS   Disclosure; Figure 1; 10pp; English.
XX
CC   AAW17083-W17088 are peptides representing mutant versions of amino acids
CC   130-163 of wild-type foot and mouth disease (FMD) virus VP-1 structural
CC   protein. The mutants were made in order to investigate the requirements
CC   of this region for cell binding. The sequences are mutant versions of
CC   the G-H loop, a flexible loop between the G and H beta strands of VP-1)
CC   VP-1, which is the main antigenic site of VP-1 and is also involved in
CC   cell binding. The FMD virus G-H loop peptide can be modified in a FMD
CC   virus by deletion of the amino acid (aa) sequence
CC   Gly-Val-Arg-Gly-Asp-Phe (aa 142-147) and replacing this sequence with the
CC   sequence Asn-Pro. This results in the virus retaining its immunogenicity
CC   but is not infectious because the cell binding site has been removed. The
CC   genetically modified FMD virus is useful in vaccines against the
CC   disease.
XX
SQ   Sequence   31 AA;

  Query Match             86.9%;  Score 86;  DB 18;  Length 31;
  Best Local Similarity   94.7%;  Pred. No. 6.7e-06;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||||| ||||||||||
Db         11 GSGVRGDKGSLALRVARQL 29


RESULT 4
ABU08643
ID   ABU08643 standard; peptide; 19 AA.
XX
AC   ABU08643;
XX
DT   04-JUN-2003  (first entry)
XX
DE   Foot and mouth disease virus (FMDV) VP1-A12141-159, variant FL.
XX
KW   Foot and mouth disease virus; FMDV; VP1-A12 region; immunosuppressive;
KW   antiinflammatory; antiarthritic; antirheumatic; dermatological;
KW   antidiabetic; antianaemic; neuroprotective; vaccine;
KW   anti-metallothionein antibody; humoral immune response;
KW   autoimmune disease; Grave's disease; rheumatoid arthritis;
KW   systemic lupus erythematosus; Type 1 diabetes; pernicious anaemia;
KW   multiple sclerosis; Sjorgen's disease.
XX
OS   Foot and mouth disease virus.
XX
PN   US2003007973-A1.
XX
PD   09-JAN-2003.
XX
PF   24-JUN-2002; 2002US-0178909.
XX
PR   22-JUN-2001; 2001US-300346P.
XX
PA   (LYNE/) LYNES M A.
XX
PI   Lynes MA;
XX
DR   WPI; 2003-353082/33.
XX
PT   Treating autoimmune diseases, e.g. Grave's disease, arthritis, lupus,
PT   or diabetes by administering a composition consisting of an
PT   anti-metallothionein antibody to stimulate a humoral immune response in
PT   a subject -
XX
PS   Example 7; Fig 9; 24pp; English.
XX
CC   The invention describes a method of treating a subject comprising
CC   administering to the subject a composition consisting of an
CC   anti-metallothionein antibody to stimulate a humoral immune response in
CC   a subject. The methods and compositions are useful for treating
CC   autoimmune diseases, e.g. Grave's disease, rheumatoid arthritis,
CC   systemic lupus erythematosus, Type 1 diabetes, pernicious anaemia,
CC   multiple sclerosis or Sjorgen's disease. This is the amino acid
CC   sequence of foot and mouth disease virus (FMDV) VP1-A12141-159 region
CC   used as an immunogen.
XX
SQ   Sequence   19 AA;

  Query Match             84.8%;  Score 84;  DB 24;  Length 19;
  Best Local Similarity   94.7%;  Pred. No. 8.2e-06;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||||| ||||||||||
Db          1 GSGVRGDFGSLALRVARQL 19


RESULT 5
ABU08642
ID   ABU08642 standard; peptide; 19 AA.
XX
AC   ABU08642;
XX
DT   04-JUN-2003  (first entry)
XX
DE   Foot and mouth disease virus (FMDV) VP1-A12141-159, variant FP.
XX
KW   Foot and mouth disease virus; FMDV; VP1-A12 region; immunosuppressive;
KW   antiinflammatory; antiarthritic; antirheumatic; dermatological;
KW   antidiabetic; antianaemic; neuroprotective; vaccine;
KW   anti-metallothionein antibody; humoral immune response;
KW   autoimmune disease; Grave's disease; rheumatoid arthritis;
KW   systemic lupus erythematosus; Type 1 diabetes; pernicious anaemia;
KW   multiple sclerosis; Sjorgen's disease.
XX
OS   Foot and mouth disease virus.
XX
PN   US2003007973-A1.
XX
PD   09-JAN-2003.
XX
PF   24-JUN-2002; 2002US-0178909.
XX
PR   22-JUN-2001; 2001US-300346P.
XX
PA   (LYNE/) LYNES M A.
XX
PI   Lynes MA;
XX
DR   WPI; 2003-353082/33.
XX
PT   Treating autoimmune diseases, e.g. Grave's disease, arthritis, lupus,
PT   or diabetes by administering a composition consisting of an
PT   anti-metallothionein antibody to stimulate a humoral immune response in
PT   a subject -
XX
PS   Example 7; Fig 9; 24pp; English.
XX
CC   The invention describes a method of treating a subject comprising
CC   administering to the subject a composition consisting of an
CC   anti-metallothionein antibody to stimulate a humoral immune response in
CC   a subject. The methods and compositions are useful for treating
CC   autoimmune diseases, e.g. Grave's disease, rheumatoid arthritis,
CC   systemic lupus erythematosus, Type 1 diabetes, pernicious anaemia,
CC   multiple sclerosis or Sjorgen's disease. This is the amino acid
CC   sequence of foot and mouth disease virus (FMDV) VP1-A12141-159 region
CC   used as an immunogen.
XX
SQ   Sequence   19 AA;

  Query Match             77.8%;  Score 77;  DB 24;  Length 19;
  Best Local Similarity   89.5%;  Pred. No. 9.2e-05;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||||| |||| ||||||
Db          1 GSGVRGDFGSLAPRVARQL 19


RESULT 6
AAY94587
ID   AAY94587 standard; Peptide; 20 AA.
XX
AC   AAY94587;
XX
DT   10-JAN-2001  (first entry)
XX
DE   Envelope gene epitope of foot and mouth disease virus.
XX
KW   Hepatitis B virus nucleocapsid antigen; HBcAg; T cell epitope;
KW   cytolytic T lymphocyte; immunogenic; ICE; CTL; FMDV;
KW   immunodominant core epitope; immunisation; envelope gene.
XX
OS   Foot and mouth disease virus.
XX
```

PN WO200026385-A1.
XX
PD 11-MAY-2000.
XX
PF 05-NOV-1999; 99WO-US26291.
XX
PR 05-NOV-1998; 98US-0107169.
XX
PA (POWD-) POWDERJECT VACCINES INC.
XX
PI Fuller DL, Fuller JT;
XX
DR WPI; 2000-451623/39.
XX
PT Use of expression vector for nucleic acid immunization that comprises
PT promoter and recombinant nucleic acid sequences encoding Hepatitis B
PT core antigen and T cell epitope from antigen -
XX
PS Example 6; Page 39; 55pp; English.
XX
CC The present invention relates to an immunogeneic recombinant
CC nucleic acid molecule. The molecule consists of a modified hepatitis
CC B virus nucleocapsid antigen (HBcAg) with a T cell epitope sequence
CC inserted within the HBcAg. The creation of a unique restriction site
CC in HBcAg facilitated the insertion of the T cell epitope into the DNA
CC encoding the immunodominant core epitope of the HBcAg. An example of a
CC suitable insertion epitope is the present sequence, the
CC neutralisation epitope from the foot and mouth disease virus envelope
CC gene. Alternatively other T cell epitopes may be inserted
CC (AAY94583, AAY94584, AAY94585, AAY94586, AAY94588). The recombinant
CC nucleic acid molecule may then be used as a reagent in various nucleic
CC acid immunisation strategies. The advantage of this method of
CC immunisation is that the nucleic acid reagents that encode hybrid HBcAg
CC generate an extremely high frequency cellular immune response against
CC the CTL epitope.
XX
SQ Sequence 20 AA;

Query Match 77.8%; Score 77; DB 21; Length 20;
Best Local Similarity 89.5%; Pred. No. 9.7e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GSGVRGDSGSLALRVARQL 20
||||||| |||| ||||||
Db 1 GSGVRGDFGSLAPRVARQL 19

RESULT 7
AAP30107
ID AAP30107 standard; Peptide; 31 AA.
XX
AC AAP30107;
XX
DT 03-APR-1992 (first entry)
XX
DE Sequence of VP1 capisid protein residues 130-160 from the amino-
DE terminus, FMDV, Tübingen type A, subtype 12, strain 119.
XX
KW Antigen; Picornavirus; capsid protein; antibody; detection;
KW vaccine; diagnosis.
XX
OS Foot and mouth disease virus.
XX
PN WO8303547-A.
XX
PD 27-OCT-1983.
XX
PF 14-APR-1983; 83WO-0002644.
XX
PR 25-MAR-1983; 83US-0478847.
PR 14-APR-1982; 82US-0368308.
PR 20-SEP-1984; 84US-0653475.
PR 18-DEC-1984; 84US-0682819.
XX
PA (BITT/) BITTLE J L.
PA (SCRI-) SCRIPPS CLINIC & RE.
XX
PI Bittle JL, Lerner RA;
XX
DR WPI; 1983-807942/44.
XX
PT Antigenic peptide(s) corresp. to picornavirus capsid protein -
PT useful in prodn. of vaccines and in diagnostic tests
XX
PS Example; Page 26; 90pp; English.
XX
CC The peptides of the invention corresp. to a region on the antigenic
CC Picornavirus capsid protein. The capsid protein FMDV VP1 or polio
CC virus VP1. When linked to carriers the peptides are immunogenic.
CC Dose is 20 ug-2mg peptide for inoculations.
XX
SQ Sequence 31 AA;

Query Match 77.8%; Score 77; DB 4; Length 31;
Best Local Similarity 89.5%; Pred. No. 0.00015;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GSGVRGDSGSLALRVARQL 20
||||||| |||| ||||||
Db 11 GSGVRGDFGSLAPRVARQL 29

RESULT 8
AAP50728
ID AAP50728 standard; protein; 31 AA.
XX
AC AAP50728;
XX
DT 23-OCT-1991 (first entry)
XX
DE Foot-and-mouth disease antigenic peptide.
XX
KW Foot-and-mouth disease; vaccine; antigen;
XX
OS Foot-and-mouth disease virus.
XX
PN US4544500-A.
XX
PD 01-OCT-1985.
XX
PF 18-DEC-1984; 84US-0682819.
XX
PR 18-DEC-1984; 84US-0682819.
XX
PA (SCRI-) SCRIPPS CLINIC & RESEARCH FOUNDATION.
XX
PI Bittle JL, Lerner RA;
XX
DR WPI; 1985-262823/42.
XX
PT Synthetic foot and mouth disease antigen - comprising 20 amino acid
PT peptide corresp. to virus protein VP1.
XX
PS Disclosure; page 4; 7pp; English.
XX
CC The peptide corresponds to amino acids 130-160 of the foot and mouth
CC disease virus VP1 protein. It represents a monospecific synthetic
CC antigenic determinant. The peptide can be used to vaccinate animals,
CC esp. swine and cattle, against the virus with a single inoculation.
XX
SQ Sequence 31 AA;

Query Match 77.8%; Score 77; DB 6; Length 31;
Best Local Similarity 89.5%; Pred. No. 0.00015;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||||| |||| ||||||
Db         11 GSGVRGDFGSLAPRVARQL 29


RESULT 9
AAW17082
ID   AAW17082 standard; peptide; 31 AA.
XX
AC   AAW17082;
XX
DT   13-JUN-1997  (first entry)
XX
DE   Foot and mouth disease virus G-H loop of protein VP-1.
XX
KW   Foot and mouth disease; aphthovirus; vaccine; antigen; epitope;
KW   immunogen; G-H loop; VP-1; deletion; mutant; binding site removal;
KW   antibody; non-infectious; attenuated.
XX
OS   Foot and mouth disease virus.
XX
PN   US5612040-A.
XX
PD   18-MAR-1997.
XX
PF   07-APR-1995;   95US-0418716.
XX
PR   07-APR-1995;   95US-0418716.
XX
PA   (USDA ) US SEC OF AGRIC.
XX
PI   Baxt B,  Berinstein A,  Kang AS,  Mason PW,  Reider E;
XX
DR   WPI; 1997-192081/17.
XX
PT   Genetically modified foot-and-mouth disease virus - lacks
PT   cell-binding site and is non-infectious, useful in vaccines
XX
PS   Disclosure; Figure 1; 10pp; English.
XX
CC   AAW17082 represents amino acids 130-163 of wild-type foot and mouth
CC   disease (FMD) virus  VP-1 structural protein. The sequence represents
CC   the G-H loop, a flexible loop between the G and H beta strands of VP-1.
CC   The G-H loop is the main antigenic site of VP-1 and is also involved in
CC   cell binding. The FMD virus G-H loop peptide can be modified in a FMD
CC   virus by deletion of the amino acid (aa) sequence
CC   Gly-Val-Arg-Gly-Asp-Phe (aa 142-147) and replacing this sequence with the
CC   sequence Asn-Pro. This results in the virus retaining its immunogenicity
CC   but is not infectious because the cell binding site has been removed. The
CC   genetically modified FMD virus is useful in vaccines against the
CC   disease.
XX
SQ   Sequence   31 AA;

  Query Match             77.8%;  Score 77;  DB 18;  Length 31;
  Best Local Similarity   89.5%;  Pred. No. 0.00015;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||||| |||| ||||||
Db         11 GSGVRGDFGSLAPRVARQL 29


RESULT 10
AAY68501
ID   AAY68501 standard; peptide; 35 AA.
XX
AC   AAY68501;
XX
DT   05-MAY-2000  (first entry)
XX
DE   Target antigenic peptide derived from amino acids 134-169 of VP1.
XX
KW   Antigenic peptide; VP1 capsid protein; FMDV strain A12; epitope;
KW   helper T-cell epitope; immune response; vaccine; FMDV infection.
XX
OS   Foot and mouth disease virus.
XX
PN   WO9966954-A1.
XX
PD   29-DEC-1999.
XX
PF   21-JUN-1999;   99WO-US13921.
XX
PR   20-JUN-1998;   98US-0100600.
XX
PA   (UNBI-) UNITED BIOMEDICAL INC.
XX
PI   Wang CY,  Shen M;
XX
DR   WPI; 2000-160563/14.
XX
PT   Synthetic peptide used in protecting animals against Foot-and-Mouth
PT   Disease Virus infections comprises a  helper T-cell epitope and a
PT   Foot-and-Mouth Disease Virus epitope conjugate -
XX
PS   Claim 1; Page 45; 115pp; English.
XX
CC   The present sequence represents a target antigenic peptide derived from
CC   the VP1 capsid protein of Foot and Mouth Disease Virus (FMDV) strain A12.
CC   The peptide is used to design synthetic epitopes, which are conjugated
CC   to a helper T-cell epitope and optionally an immunostimulator sequence.
CC   The peptides are covalently linked to each other, preferably by peptide
CC   bonds generated through direct synthesis. The synthetic vaccines of the
CC   invention require lower levels of biosecurity, have a reduced loss of
CC   immunogenic activity, are less prone to product instability and
CC   lot-to-lot variability and do not need periodic revision due to
CC   antigenic variation in the field. The peptides are used as immunogens
CC   in vaccines to generate an enhanced immune response against a FMDV
CC   antigen in animals, such as swine, cattle, sheep, goats and susceptible
CC   wild species. They can also be used to diagnose FMDV infection in a
CC   mammal.
XX
SQ   Sequence   35 AA;

  Query Match             77.8%;  Score 77;  DB 21;  Length 35;
  Best Local Similarity   89.5%;  Pred. No. 0.00017;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||||| |||| ||||||
Db          7 GSGVRGDFGSLAPRVARQL 25


RESULT 11
AAY68528
ID   AAY68528 standard; peptide; 75 AA.
XX
AC   AAY68528;
XX
DT   05-MAY-2000  (first entry)
XX
DE   Synthetic foot and mouth disease virus immunogen.
XX
KW   Antigenic peptide; VP1 capsid protein; serotype Asia; epitope;
KW   helper T-cell epitope; immune response; vaccine; FMDV infection.
XX
OS   Synthetic.
OS   Yersinia sp.
OS   Foot and mouth disease virus.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..16
FT                   /note= "Yersinia invasin domain (immunostimulator
FT                         peptide"
FT   Peptide         17..18
```

```
FT                  /note= "spacer"
FT   Peptide        19..38
FT                  /note= "autologous helper T-cell epitope derived
FT                   from VP1 capsid protein"
FT   Peptide        39..40
FT                  /note= "spacer"
FT   Peptide        41..75
FT                  /note= "synthetic VP1 capsid protein epitope"
XX
PN   WO9966954-A1.
XX
PD   29-DEC-1999.
XX
PF   21-JUN-1999;   99WO-US13921.
XX
PR   20-JUN-1998;   98US-0100600.
XX
PA   (UNBI-) UNITED BIOMEDICAL INC.
XX
PI   Wang CY,  Shen M;
XX
DR   WPI; 2000-160563/14.
XX
PT   Synthetic peptide used in protecting animals against Foot-and-Mouth
PT   Disease Virus infections comprises a  helper T-cell epitope and a
PT   Foot-and-Mouth Disease Virus epitope conjugate -
XX
PS   Claim 8; Page 96; 115pp; English.
XX
CC   The present sequence represents a foot and mouth disease (FMDV) vaccine
CC   of the invention. The specification describes epitopes derived from an
CC   antigen of the VP1 capsid protein of FMDV. The VP1 epitopes are
CC   conjugated to a helper T-cell epitope and optionally an immunostimulator
CC   sequence. The peptides are covalently linked to each other, preferably
CC   by peptide bonds generated through direct synthesis. The synthetic
CC   vaccines of the invention require lower levels of biosecurity, have a
CC   reduced loss of immunogenic activity, are less prone to product
CC   instability and lot-to-lot variability and do not need periodic revision
CC   due to antigenic variation in the field. The peptides are used as
CC   immunogens in vaccines to generate an enhanced immune response against
CC   a FMDV antigen in animals, such as swine, cattle, sheep, goats and
CC   susceptible wild species. They can also be used to diagnose FMDV
CC   infection in a mammal.
XX
SQ   Sequence   75 AA;

  Query Match             77.8%;  Score 77;  DB 21;  Length 75;
  Best Local Similarity   89.5%;  Pred. No. 0.00037;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||||| |||| ||||||
Db         47 GSGVRGDFGSLAPRVARQL 65


RESULT 12
AAP30205
ID   AAP30205 standard; Protein; 216 AA.
XX
AC   AAP30205;
XX
DT   29-JUL-1992  (first entry)
XX
DE   Sequence encoded by direct expression vector pFM 10 which
DE   comprises the promoter and operator of the E. coli tryptophan
DE   operon AAs 1-211 of the VP3 gene linked to 4 AAs from pBR322.
XX
KW   Vaccine; immunogen; antigen; viral protein.
XX
OS   Foot and mouth disease virus.
XX
PN   EP68693-A.
XX
PD   05-JAN-1983.
XX
PF   11-JUN-1982;   82EP-0303040.
XX
PR   04-MAY-1982;   82US-0374855.
PR   16-JUN-1981;   81US-0274103.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Kleid DG,  Yansura DG;
XX
DR   WPI; 1983-05055K/03.
DR   N-PSDB; AAN30144.
XX
PT   Polypeptide antigens of foot-and-mouth disease - obtd. by
PT   recombinant DNA technology for vaccine prodn.
XX
PS   Example; Page 36-37; 81pp; English.
XX
CC   The inventors claim a polypeptide comprising at least one antigen of
CC   FMDV, or a fusion polypeptide of at least two antigens from at least
CC   two different FMDV strands, and encoding DNA.  Pref. the fusion
CC   polypeptide has the sequence of two of the A, C, O, Asia-1, SAT 1,
CC   SAT 2 and SAT 3 type strains of the virus.  The polypeptides pref.
CC   comprise AAs 55-177 or 1-210 of the VP3 virus.  Fusion polypeptides
CC   may comprise AAs 1-210 fused to 8-210 of VP3 virus protein.  pFM1,
CC   pFM2, pFM3, pFM10, pFM20, pFMB1, pFMC, pFMD, pFMF and pFMG are also
CC   claimed.
XX
SQ   Sequence   216 AA;

  Query Match             77.8%;  Score 77;  DB 4;  Length 216;
  Best Local Similarity   89.5%;  Pred. No. 0.0011;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||||| |||| ||||||
Db        141 GSGVRGDFGSLAPRVARQL 159


RESULT 13
AAP30195
ID   AAP30195 standard; Protein; 220 AA.
XX
AC   AAP30195;
XX
DT   29-JUL-1992  (first entry)
XX
DE   Sequence encoded by VP3 gene of FMDV type A12 (T465).
XX
KW   Vaccine; immunogen; antigen; viral protein.
XX
OS   Foot and mouth disease virus.
XX
PN   EP68693-A.
XX
PD   05-JAN-1983.
XX
PF   11-JUN-1982;   82EP-0303040.
XX
PR   04-MAY-1982;   82US-0374855.
PR   16-JUN-1981;   81US-0274103.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Kleid DG,  Yansura DG;
XX
DR   WPI; 1983-05055K/03.
DR   N-PSDB; AAN30127.
XX
PT   Polypeptide antigens of foot-and-mouth disease - obtd. by
PT   recombinant DNA technology for vaccine prodn.
XX
```

```
PS   Disclosure; Fig 2; 81pp; English.
XX
CC   The inventors claim a polypeptide comprising at least one antigen of
CC   FMDV, or a fusion polypeptide of at least two antigens from at least
CC   two different FMDV strands, and encoding DNA.  Pref. the fusion
CC   polypeptide has the sequence of two of the A, C, O, Asia-1, SAT 1,
CC   SAT 2 and SAT 3 type strains of the virus.  The polypeptides pref.
CC   comprise AAs 55-177 or 1-210 of the VP3 virus.  Fusion polypeptides
CC   may comprise AAs 1-210 fused to 8-210 of VP3 virus protein.  pFM1,
CC   pFM2, pFM3, pFM10, pFM20, pFMB1, pFMC, pFMD, pFMF and pFMG are also
CC   claimed.
XX
SQ   Sequence   220 AA;

  Query Match            77.8%;  Score 77;  DB 4;  Length 220;
  Best Local Similarity  89.5%;  Pred. No. 0.0011;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||||| |||| ||||||
Db        140 GSGVRGDFGSLAPRVARQL 158


RESULT 14
AAP30204
ID   AAP30204 standard; Protein; 233 AA.
XX
AC   AAP30204;
XX
DT   29-JUL-1992  (first entry)
XX
DE   Sequence encoded by pFM3 which comprises 17 AAs from the LE' gene
DE   construction (fragment 3a), linked to a methionine AA (CNBr cleavable),
DE   linked to the AAs 1-211 of the VP3 FMDV A12, linked to 4 AAs from
DE   pBR322.
XX
KW   Vaccine; immunogen; antigen; viral protein.
XX
OS   Foot and mouth disease virus and Escherichia coli.
XX
PN   EP68693-A.
XX
PD   05-JAN-1983.
XX
PF   11-JUN-1982;   82EP-0303040.
XX
PR   04-MAY-1982;   82US-0374855.
PR   16-JUN-1981;   81US-0274103.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Kleid DG,  Yansura DG;
XX
DR   WPI; 1983-05055K/03.
DR   N-PSDB; AAN30142.
XX
PT   Polypeptide antigens of foot-and-mouth disease - obtd. by
PT   recombinant DNA technology for vaccine prodn.
XX
PS   Example; Page 33-34; 81pp; English.
XX
CC   The inventors claim a polypeptide comprising at least one antigen of
CC   FMDV, or a fusion polypeptide of at least two antigens from at least
CC   two different FMDV strands, and encoding DNA.  Pref. the fusion
CC   polypeptide has the sequence of two of the A, C, O, Asia-1, SAT 1,
CC   SAT 2 and SAT 3 type strains of the virus.  The polypeptides pref.
CC   comprise AAs 55-177 or 1-210 of the VP3 virus.  Fusion polypeptides
CC   may comprise AAs 1-210 fused to 8-210 of VP3 virus protein.  pFM1,
CC   pFM2, pFM3, pFM10, pFM20, pFMB1, pFMC, pFMD, pFMF and pFMG are also
CC   claimed.
XX
SQ   Sequence   233 AA;

  Query Match            77.8%;  Score 77;  DB 4;  Length 233;
  Best Local Similarity  89.5%;  Pred. No. 0.0012;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||||| |||| ||||||
Db        158 GSGVRGDFGSLAPRVARQL 176


RESULT 15
AAP30202
ID   AAP30202 standard; Protein; 405 AA.
XX
AC   AAP30202;
XX
DT   29-JUL-1992  (first entry)
XX
DE   Sequence encoded by pFM1 which comprises 190 AAs form the LE' gene
DE   construction linked to 6 AAs from the linker molecule (fragment 4),
DE   linked to 205 AAs from the VP3 FMDV A12 (codons 7-211), linked to
DE   4 AAs from pBR322.
XX
KW   Vaccine; immunogen; antigen; viral protein.
XX
OS   Foot and mouth disease virus and Escherichia coli.
XX
PN   EP68693-A.
XX
PD   05-JAN-1983.
XX
PF   11-JUN-1982;   82EP-0303040.
XX
PR   04-MAY-1982;   82US-0374855.
PR   16-JUN-1981;   81US-0274103.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Kleid DG,  Yansura DG;
XX
DR   WPI; 1983-05055K/03.
DR   N-PSDB; AAN30140.
XX
PT   Polypeptide antigens of foot-and-mouth disease - obtd. by
PT   recombinant DNA technology for vaccine prodn.
XX
PS   Example; Page 29-31; 81pp; English.
XX
CC   The inventors claim a polypeptide comprising at least one antigen of
CC   FMDV, or a fusion polypeptide of at least two antigens from at least
CC   two different FMDV strands, and encoding DNA.  Pref. the fusion
CC   polypeptide has the sequence of two of the A, C, O, Asia-1, SAT 1,
CC   SAT 2 and SAT 3 type strains of the virus.  The polypeptides pref.
CC   comprise AAs 55-177 or 1-210 of the VP3 virus.  Fusion polypeptides
CC   may comprise AAs 1-210 fused to 8-210 of VP3 virus protein.  pFM1,
CC   pFM2, pFM3, pFM10, pFM20, pFMB1, pFMC, pFMD, pFMF and pFMG are also
CC   claimed.
XX
SQ   Sequence   405 AA;

  Query Match            77.8%;  Score 77;  DB 4;  Length 405;
  Best Local Similarity  89.5%;  Pred. No. 0.002;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||||| |||| ||||||
Db        330 GSGVRGDFGSLAPRVARQL 348


Search completed: October  9, 2003, 10:14:18
Job time : 58.3333 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:10:09 ; Search time 19.6667 Seconds
(without alignments)
97.799 Million cell updates/sec

Title: US-09-549-186B-9
Perfect score: 99
Sequence: 1 CGSGVRGDSGSLALRVARQL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	77	77.8	2332	1	GNNY4F	genome polyprotein
2	66	66.7	216	2	A03911	genome polyprotein
3	50	50.5	76	2	D82844	carbon storage reg
4	50	50.5	170	2	T35564	ribosomal protein
5	49	49.5	230	2	A03909	genome polyprotein
6	48	48.5	2336	2	S37077	genome polyprotein
7	46	46.5	134	2	F72250	ribosomal protein
8	46	46.5	312	2	A56911	TRADD protein - hu
9	46	46.5	2333	1	GNNY2F	genome polyprotein
10	45	45.5	398	2	A58922	esterase/N-deacety
11	45	45.5	467	2	B75322	probable oligoendo
12	45	45.5	512	2	S06115	alpha-amylase (EC
13	44	44.4	109	2	S61252	genome polyprotein
14	44	44.4	109	2	S61253	genome polyprotein
15	44	44.4	136	2	H90573	ribosomal protein
16	44	44.4	537	2	AB1869	type I site-specif
17	43	43.4	122	2	S72797	probable K+ transp
18	43	43.4	130	2	E82669	30S ribosomal prot
19	43	43.4	151	2	H70976	probable ribosomal
20	43	43.4	153	2	S72982	ribosomal protein
21	43	43.4	211	2	T35272	hypothetical prote
22	43	43.4	281	1	S34626	translation elonga
23	43	43.4	303	2	E72463	hypothetical prote
24	43	43.4	372	2	AI1573	NADPH-dependent bu
25	43	43.4	439	2	S73586	probable lipoprote
26	42	42.4	99	2	F82709	hypothetical prote
27	42	42.4	239	2	A61167	corticotropin / li
28	42	42.4	250	2	H87257	hypothetical prote
29	42	42.4	326	2	A99757	homolog of Salmone
30	42	42.4	326	2	G85620	homolog of Salmone
31	42	42.4	356	2	D64834	probable membrane
32	42	42.4	372	2	AF1220	NADPH-dependent bu
33	42	42.4	373	2	T50605	hypothetical prote
34	42	42.4	439	2	A45068	5'-3' DNA helicase
35	42	42.4	500	1	S60929	probable aldehyde
36	42	42.4	694	2	T33561	hypothetical prote
37	42	42.4	809	1	S60112	transitional endop
38	42	42.4	1015	1	JS0628	formate dehydrogen
39	42	42.4	1015	2	C85729	hypothetical prote
40	42	42.4	1015	2	F90888	hypothetical prote
41	42	42.4	1062	2	D96540	hypothetical prote
42	42	42.4	1159	2	AH3088	IcmF [imported] -
43	42	42.4	1159	2	B98198	hypothetical prote
44	42	42.4	1318	2	T39066	hypothetical prote
45	42	42.4	1446	2	T13018	hypothetical prote

ALIGNMENTS

RESULT 1
GNNY4F
genome polyprotein - foot-and-mouth disease virus A (strain A12)
N;Contains: coat protein VP1; coat protein VP2; coat protein VP3; coat protein VP4; c
tein VPg2; genome-linked protein VPg3; nonstructural protein p20a; proteinase (EC 3.4
C;Species: Aphthovirus A (foot-and-mouth disease virus A)
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C;Accession: A25794
R;Robertson, B.H.; Grubman, M.J.; Weddell, G.N.; Moore, D.M.; Welsh, J.D.; Fischer, T
J. Virol. 54, 651-660, 1985
A;Title: Nucleotide and amino acid sequence coding for polypeptides of foot-and-mouth
A;Reference number: A25794; MUID:85211015; PMID:2987518
A;Accession: A25794
A;Molecule type: genomic RNA
A;Residues: 1-2332 <ROB>
A;Cross-references: GB:M10975; NID:g210306; PIDN:AAA42593.1; PID:g210307
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: coat protein; core protein; genome-linked protein; hydrolase; nonstructur
F;1-216/Product: nonstructural protein p20a #status predicted <NPA>
F;217-285/Product: coat protein VP4 #status predicted <VP4>
F;286-503/Product: coat protein VP2 #status predicted <VP2>
F;504-723/Product: coat protein VP3 #status predicted <VP3>
F;724-937/Product: coat protein VP1 #status predicted <VP1>
F;938-953/Product: core protein X #status predicted <CPX>
F;954-1107/Product: core protein p14 #status predicted <C14>
F;1108-1425/Product: core protein p41 #status predicted <C41>
F;1426-1578/Product: core protein p19 #status predicted <C19>
F;1579-1601/Product: genome-linked protein VPg1 #status predicted <VG1>
F;1602-1625/Product: genome-linked protein VPg2 #status predicted <VG2>
F;1626-1649/Product: genome-linked protein VPg3 #status predicted <VG3>
F;1650-1862/Product: proteinase #status predicted <PTS>
F;1863-2332/Product: RNA-directed RNA polymerase #status predicted <RRP>

Query Match 77.8%; Score 77; DB 1; Length 2332;
Best Local Similarity 89.5%; Pred. No. 0.002;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||||| |||| ||||||
Db        864 GSGVRGDFGSLAPRVARQL 882
```

RESULT 2
A03911
genome polyprotein - foot-and-mouth disease virus A (strain A24 Cruzeiro) (fragment)
N;Contains: coat protein VP1; core protein p52
C;Species: Aphthovirus A (foot-and-mouth disease virus A)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 26-Aug-1999
C;Accession: A03911
R;Makoff, A.J.; Paynter, C.A.; Rowlands, D.J.; Boothroyd, J.C.
Nucleic Acids Res. 10, 8285-8295, 1982
A;Title: Comparison of the amino acid sequence of the major immunogen from three sero

```
A;Reference number: A37503; MUID:83143292; PMID:6298715
A;Accession: A03911
A;Molecule type: genomic RNA
A;Residues: 1-216 <MAK>
A;Cross-references: GB:J02183; NID:g210312; PIDN:AAA42596.1; PID:g210313
C;Comment: Specific enzymatic cleavage in vivo yields mature proteins including coat prc
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: polyprotein

  Query Match             66.7%;  Score 66;  DB 2;  Length 216;
  Best Local Similarity   73.7%;  Pred. No. 0.0089;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||| ||| |||| || :||
Db        129 GSGRRGDMGSLAARVVKQL 147


RESULT 3
D82844
carbon storage regulator XF0125 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 17-Nov-2000
C;Accession: D82844
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: D82844
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-76 <SIM>
A;Cross-references: GB:AE003866; GB:AE003849; NID:g9104906; PIDN:AAF82938.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0125
C;Superfamily: glycogen biosynthesis inhibitor

  Query Match             50.5%;  Score 50;  DB 2;  Length 76;
  Best Local Similarity   81.8%;  Pred. No. 0.94;
  Matches    9;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDSGS 11
              | || ||||||
Db         64 CSSGTRGDSGS 74


RESULT 4
T35564
ribosomal protein S9 - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 19-May-2000
C;Accession: T35564
R;Saunders, D.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1998
A;Reference number: Z21583
A;Accession: T35564
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-170 <SAU>
A;Cross-references: EMBL:AL031317; PIDN:CAA20391.1; GSPDB:GN00070; SCOEDB:SC6G4.13
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: rpsI; SCOEDB:SC6G4.13
C;Superfamily: Escherichia coli ribosomal protein S9

  Query Match             50.5%;  Score 50;  DB 2;  Length 170;
  Best Local Similarity   57.9%;  Pred. No. 2.1;
  Matches   11;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              | || | :|:| | ||| |
Db        108 GGGVSGQAGALRLGVARAL 126


RESULT 5
A03909
genome polyprotein - foot-and-mouth disease virus A (strain A5) (fragment)
N;Contains: coat protein VP1; coat protein VP3; core protein p52
C;Species: Aphthovirus A (foot-and-mouth disease virus A)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 07-Feb-1997
C;Accession: A03909
R;Beck, E.; Feil, G.; Strohmaier, K.
EMBO J. 2, 555-559, 1983
A;Title: The molecular basis of the antigenic variation of foot-and-mouth disease vir
A;Reference number: A03909; MUID:84028562; PMID:6194987
A;Accession: A03909
A;Molecule type: mRNA
A;Residues: 1-230 <BEC>
A;Note: the authors translated the codon GCA for residue 114 as Val
C;Comment: Specific enzymatic cleavages in vivo yield mature proteins including the c
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: polyprotein

  Query Match             49.5%;  Score 49;  DB 2;  Length 230;
  Best Local Similarity   57.9%;  Pred. No. 4.1;
  Matches   11;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              |   ||| || | | |:||
Db        143 GGPRRGDMGSAAARAAKQL 161


RESULT 6
S37077
genome polyprotein - foot-and-mouth disease virus A (strain A22/550 Azerbaijan 65)
N;Contains: coat protein VP1; coat protein VP2; coat protein VP3; coat protein VP4; c
tein VPg2; genome-linked protein VPg3; nonstructural protein p20a; proteinase (EC 3.4
C;Species: Aphthovirus A (foot-and-mouth disease virus A)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 26-Aug-1999
C;Accession: S37077; JN0413
R;Sosnovtsev, S.V.; Onischenko, A.M.; Petrov, N.A.; Kalashnikova, T.I.; Mamaeva, N.V.
submitted to the EMBL Data Library, August 1993
A;Reference number: S37077
A;Accession: S37077
A;Molecule type: genomic RNA
A;Residues: 1-2336 <SOS>
A;Cross-references: EMBL:X74812; NID:g397965; PIDN:CAA52812.1; PID:g397966
R;Onishchenko, A.M.; Petrov, N.A.; Blinov, V.M.; Vassilenko, S.K.; Sandakhchiev, L.S.
Bioorg. Khim. 12, 416-419, 1986
A;Title: The DNA-copy primary structure for the gene of the FMDV A22 serotype VP1-pro
A;Reference number: JN0413; MUID:86186952; PMID:2421736
A;Accession: JN0413
A;Molecule type: genomic RNA
A;Residues: 702-955 <ONI>
A;Cross-references: GB:M38362; NID:g210514; PIDN:AAA42664.1; PID:g210515
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: coat protein; core protein; genome-linked protein; hydrolase; nonstructur
F;1-217/Product: nonstructural protein p20a #status predicted <NPA>
F;218-286/Product: coat protein VP4 #status predicted <VP4>
F;287-504/Product: coat protein VP2 #status predicted <VP2>
```

```
F;505-724/Product: coat protein VP3 #status predicted <VP3>
F;725-938/Product: coat protein VP1 #status predicted <VP1>
F;939-954/Product: core protein X #status predicted <CPX>
F;955-1108/Product: core protein p14 #status predicted <C14>
F;1109-1426/Product: core protein p41 #status predicted <C41>
F;1427-1579/Product: core protein p19 #status predicted <C19>
F;1580-1602/Product: genome-linked protein VPg1 #status predicted <VG1>
F;1603-1626/Product: genome-linked protein VPg2 #status predicted <VG2>
F;1627-1650/Product: genome-linked protein VPg3 #status predicted <VG3>
F;1651-1863/Product: proteinase #status predicted <PTS>
F;1864-2333/Product: RNA-directed RNA polymerase #status predicted <RRP>

  Query Match             48.5%;  Score 48;  DB 2;  Length 2336;
  Best Local Similarity   63.2%;  Pred. No. 61;
  Matches   12;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              | | |||   || ||| ||
Db        865 GMGRRGDLEPLAARVAAQL 883


RESULT 7
F72250
ribosomal protein S9 - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: F72250
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
 C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: F72250
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-134 <ARN>
A;Cross-references: GB:AE001797; GB:AE000512; NID:g4982017; PIDN:AAD36521.1; PID:g498201
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1453
C;Superfamily: Escherichia coli ribosomal protein S9

  Query Match             46.5%;  Score 46;  DB 2;  Length 134;
  Best Local Similarity   47.4%;  Pred. No. 6.9;
  Matches    9;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              | |: | ||:: | :|| |
Db         73 GGGLSGQSGAVRLGIARAL 91


RESULT 8
A56911
TRADD protein - human
N;Alternate names: TNF receptor 1-associated protein TRADD
C;Species: Homo sapiens (man)
C;Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 17-Mar-1999
C;Accession: A56911
R;Hsu, H.; Xiong, J.; Goeddel, D.V.
Cell 81, 495-504, 1995
A;Title: The TNF receptor 1-associated protein TRADD signals cell death and NF-kappaB ac
A;Reference number: A56911; MUID:95277836; PMID:7758105
A;Accession: A56911
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-312 <HSU>
A;Cross-references: GB:L41690
C;Genetics:
A;Gene: GDB:TRADD
A;Cross-references: GDB:1320395
C;Keywords: apoptosis

  Query Match             46.5%;  Score 46;  DB 2;  Length 312;
  Best Local Similarity   61.5%;  Pred. No. 16;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDSGSLA 13
              |||| ||  | :|
Db        164 CGSGARGGDGEVA 176


RESULT 9
GNNY2F
genome polyprotein - foot-and-mouth disease virus A (strain A[10]61)
N;Contains: coat protein VP1; coat protein VP2; coat protein VP3; coat protein VP4; c
 protein p20a; nonstructural protein p20b; RNA-directed RNA polymerase (EC 2.7.7.48)
C;Species: Aphthovirus A (foot-and-mouth disease virus A)
C;Date: 17-Dec-1982 #sequence_revision 28-Aug-1985 #text_change 16-Jul-1999
C;Accession: A93508; A91491; S30753
R;Carroll, A.R.; Rowlands, D.J.; Clarke, B.E.
Nucleic Acids Res. 12, 2461-2472, 1984
A;Title: The complete nucleotide sequence of the RNA coding for the primary translati
A;Reference number: A93508; MUID:84169547; PMID:6324120
A;Accession: A93508
A;Molecule type: genomic RNA
A;Residues: 1-2333 <CAR>
A;Cross-references: GB:X00429
R;Boothroyd, J.C.; Harris, T.J.R.; Rowlands, D.J.; Lowe, P.A.
Gene 17, 153-161, 1982
A;Title: The nucleotide sequence of cDNA coding for the structural proteins of foot-a
A;Reference number: A91491; MUID:82211814; PMID:6282711
A;Accession: A91491
A;Molecule type: genomic RNA
A;Residues: 115-395,'C',397-631,'L',633-1048 <BOO>
A;Cross-references: GB:V01130; NID:g61048; PIDN:CAA24361.1; PID:g1335402
R;Sangar, D.V.; Newton, S.E.; Rowlands, D.J.; Clarke, B.E.
Nucleic Acids Res. 15, 3305-3315, 1987
A;Title: All foot and mouth disease virus serotypes initiate protein synthesis at two
A;Reference number: S30753; MUID:87203363; PMID:3033601
A;Accession: S30753
A;Molecule type: genomic RNA
A;Residues: 1-32 <SAN>
A;Cross-references: EMBL:M31575; NID:g210486; PIDN:AAA42655.1; PID:g210487
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: coat protein; core protein; genome-linked protein; nonstructural protein;
F;1-204/Product: nonstructural protein p20a #status predicted <NPA>
F;205-286/Product: coat protein VP4 #status predicted <VP4>
F;287-504/Product: coat protein VP2 #status predicted <VP2>
F;505-725/Product: coat protein VP3 #status predicted <VP3>
F;726-937/Product: coat protein VP1 #status predicted <VP1>
F;938-1578/Product: core protein p52 #status predicted <CPP>
F;1579-1601/Product: genome-linked protein VPg1 #status predicted <GL1>
F;1602-1625/Product: genome-linked protein VPg2 #status predicted <GL2>
F;1626-1649/Product: genome-linked protein VPg3 #status predicted <GL3>
F;1650-1863/Product: nonstructural protein p20b #status predicted <NPB>
F;1864-2333/Product: RNA-directed RNA polymerase #status predicted <RRP>

  Query Match             46.5%;  Score 46;  DB 1;  Length 2333;
  Best Local Similarity   71.4%;  Pred. No. 1.2e+02;
  Matches   10;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 GDSGSLALRVARQL 20
              || ||:| ||| ||
Db        869 GDLGSIAARVATQL 882


RESULT 10
A58922
esterase/N-deacetylase (EC 3.5.1.-), 50K hepatic - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 05-Dec-1998
C;Accession: A58922
R;Ozols, J.
submitted to the Protein Sequence Database, December 1998
A;Description: Determination of Lumenal Orientation of Microsomal 50-kDa Esterase/N-D
```

```
A;Reference number: A58922
A;Accession: A58922
A;Molecule type: protein
A;Residues: 1-398 <OZO>
A;Experimental source: Oryctolagus cuniculus GV
C;Keywords: hydrolase

  Query Match             45.5%;  Score 45;  DB 2;  Length 398;
  Best Local Similarity   63.2%;  Pred. No. 30;
  Matches   12;  Conservative    2;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          4 GVRGDS--GSLALRVARQL 20
              || |||  |:||  ||:||
Db        183 GVSGDSAGGNLAAAVAQQL 201


RESULT 11
B75322
probable oligoendopeptidase F - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C;Accession: B75322
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: B75322
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-467 <WHI>
A;Cross-references: GB:AE002041; GB:AE000513; NID:g6459838; PIDN:AAF11604.1; PID:g645984
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR2055
A;Map position: 1

  Query Match             45.5%;  Score 45;  DB 2;  Length 467;
  Best Local Similarity   53.3%;  Pred. No. 35;
  Matches    8;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDSGSLALR 15
              | :||:|: |:|| |
Db        103 CMNGVKGEEGTLARR 117


RESULT 12
S06115
alpha-amylase (EC 3.2.1.1) AMY1 precursor - yeast (Schwanniomyces occidentalis)
C;Species: Schwanniomyces occidentalis
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 22-Jun-1999
C;Accession: S06115
R;Strasser, A.W.M.; Selk, R.; Dohmen, R.J.; Niermann, T.; Bielefeld, M.; Seeboth, P.; Tu
Eur. J. Biochem. 184, 699-706, 1989
A;Title: Analysis of the alpha-amylase gene of Schwanniomyces occidentalis and the secre
A;Reference number: S06115; MUID:90032659; PMID:2806251
A;Accession: S06115
A;Molecule type: DNA
A;Residues: 1-512 <STR>
A;Cross-references: EMBL:X16040; NID:g4882; PIDN:CAA34162.1; PID:g4883
C;Genetics:
A;Gene: AMY1
C;Function:
A;Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A;Pathway: glycogen/starch degradation
C;Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C;Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-512/Product: alpha-amylase #status predicted <MAT>
F;209-336/Domain: alpha-amylase core homology <AMY>

  Query Match             45.5%;  Score 45;  DB 2;  Length 512;
  Best Local Similarity   47.1%;  Pred. No. 38;
  Matches    8;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDSGSLALRVA 17
              | | : ||||||:: ::
Db        475 CNSVLAGDSGSLSVSIS 491


RESULT 13
S61252
genome polyprotein - foot-and-mouth disease virus Asia (isolate Asia I Mandya, Karnat
N;Alternate names: immunogenic polyprotein
C;Species: Aphthovirus Asia (foot-and-mouth disease virus Asia)
A;Variety: isolate Asia I Mandya, Karnataka
C;Date: 18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change 26-Aug-1999
C;Accession: S61252
R;Tulasiram, P.; Tyagi, M.; Suryanarayana, V.
submitted to the EMBL Data Library, June 1995
A;Description: Antigenic variation in foot and mouth disease virus type Asia I isolat
A;Reference number: S61252
A;Accession: S61252
A;Molecule type: mRNA
A;Residues: 1-109 <TUL>
A;Cross-references: EMBL:X88856; NID:g971407; PIDN:CAA61326.1; PID:g971408
A;Experimental source: isolate Asia I Mandya, Karnataka
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: coat protein; polyprotein; proteinase
F;1-91/Product: coat protein VP1 (fragment) #status predicted <MAT1>
F;92-107/Product: proteinase 2A #status predicted <MAT2>
F;108-109/Product: proteinase 2B (fragment) #status predicted <MAT3>

  Query Match             44.4%;  Score 44;  DB 2;  Length 109;
  Best Local Similarity   60.0%;  Pred. No. 11;
  Matches    9;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 RGDSGSLALRVARQL 20
              |||  :|| |::|||
Db         23 RGDMAALAQRLSRQL 37


RESULT 14
S61253
genome polyprotein - foot-and-mouth disease virus Asia (isolate Asia I Nilgiri, Tamil
N;Alternate names: immunogenic polyprotein
N;Contains: coat protein VP1; proteinase 2A; proteinase 2B
C;Species: Aphthovirus Asia (foot-and-mouth disease virus Asia)
A;Variety: isolate Asia I Nilgiri, Tamil Nadu
C;Date: 18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change 26-Aug-1999
C;Accession: S61253
R;Tulasiram, P.; Tyagi, M.; Suryanarayana, V.
submitted to the EMBL Data Library, June 1995
A;Description: Antigenic variation in foot and mouth disease virus type Asia I isolat
A;Reference number: S61253
A;Accession: S61253
A;Molecule type: mRNA
A;Residues: 1-109 <TUL>
A;Cross-references: EMBL:X88857; NID:g971409; PIDN:CAA61327.1; PID:g971410
A;Experimental source: isolate Asia I Nilgiri, Tamil Nadu
C;Superfamily: foot-and-mouth disease virus genome polyprotein
C;Keywords: coat protein; polyprotein; proteinase
F;1-91/Product: coat protein VP1 (fragment) #status predicted <MAT1>
F;92-107/Product: proteinase 2A #status predicted <MAT2>
F;108-109/Product: proteinase 2B (fragment) #status predicted <MAT3>

  Query Match             44.4%;  Score 44;  DB 2;  Length 109;
  Best Local Similarity   60.0%;  Pred. No. 11;
  Matches    9;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 RGDSGSLALRVARQL 20
              |||  :|| |::|||
Db         23 RGDMAALAQRLSRQL 37
```

```
RESULT 15
H90573
ribosomal protein S9 [similarity] - Mycoplasma pulmonis (strain UAB CTIP)
C;Species: Mycoplasma pulmonis
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-May-2002
C;Accession: H90573
R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A;Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A;Reference number: A99512; MUID:21267165; PMID:11353084
A;Accession: H90573
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-136 <KUR>
A;Cross-references: GB:AL445566; PID:g14089910; PIDN:CAC13669.1; GSPDB:GN00153
A;Experimental source: strain UAB CTIP
C;Genetics:
A;Gene: MYPU_4960
A;Genetic code: SGC3
C;Superfamily: Escherichia coli ribosomal protein S9

  Query Match             44.4%;  Score 44;  DB 2;  Length 136;
  Best Local Similarity   42.1%;  Pred. No. 14;
  Matches    8;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              | |: | :|:: | :|| |
Db         75 GGGLAGQAGAIRLGIARAL 93


Search completed: October  9, 2003, 10:18:44
Job time : 21.6667 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:02:54 ; Search time 10.3333 Seconds
(without alignments)
91.019 Million cell updates/sec

Title: US-09-549-186B-9
Perfect score: 99
Sequence: 1 CGSGVRGDSGSLALRVARQL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	77	77.8	2332	1	POLG_FMDVA	P03308 f genome po
2	66	66.7	216	1	POLG_FMDVC	P03309 foot-and-mo
3	50	50.5	170	1	RS9_STRCO	Q53875 streptomyce
4	49	49.5	129	1	RS9_CHLTE	Q8kbk5 chlorobium
5	49	49.5	230	1	POLG_FMDV5	P03307 foot-and-mo
6	48	48.5	2336	1	POLG_FMDVZ	P49303 f genome po
7	46	46.5	134	1	RS9_THEMA	Q9x1g4 thermotoga
8	46	46.5	312	1	TRAD_HUMAN	Q15628 homo sapien
9	46	46.5	2333	1	POLG_FMDV1	P03306 f genome po
10	45	45.5	501	1	AMPA_BUCBP	Q89ag2 buchnera ap
11	45	45.5	512	1	AMY1_DEBOC	P19269 debaryomyce
12	44	44.4	94	1	CH11_RHOSH	P25969 rhodobacter
13	44	44.4	95	1	CH10_RHOPA	Q93mh2 rhodopseudo
14	44	44.4	128	1	RR9_CYACA	Q9tlv4 cyanidium c
15	44	44.4	136	1	RS9_MYCPU	Q98q72 mycoplasma
16	44	44.4	833	1	SRC2_MOUSE	P59222 mus musculu
17	43	43.4	110	1	INS_PSAOB	Q62587 psammomys o
18	43	43.4	130	1	RS9_XANAC	Q8pq41 xanthomonas
19	43	43.4	130	1	RS9_XANCP	Q8pd66 xanthomonas
20	43	43.4	130	1	RS9_XYLFA	Q9pd43 xylella fas
21	43	43.4	151	1	RS9_MYCTU	O06259 mycobacteri
22	43	43.4	153	1	RS9_MYCLE	P40828 mycobacteri
23	43	43.4	281	1	EF1D_HUMAN	P29692 homo sapien
24	43	43.4	348	1	ADH1_PICST	O00097 pichia stip
25	43	43.4	439	1	YF82_MYCPN	P75198 mycoplasma
26	43	43.4	516	1	Y4NM_RHISN	P55585 rhizobium s
27	42	42.4	356	1	YCBT_ECOLI	P75858 escherichia
28	42	42.4	439	1	DDA_BPT4	P32270 bacteriopha
29	42	42.4	499	1	DHA6_YEAST	P54115 saccharomyc
30	42	42.4	809	1	C48A_ARATH	P54609 arabidopsis
31	42	42.4	1015	1	FDNG_ECOLI	P24183 escherichia
32	42	42.4	1318	1	YDH4_SCHPO	Q92349 schizosacch
33	42	42.4	1355	1	CA21_RANCA	O42350 rana catesb
34	41	41.4	108	1	INS1_MOUSE	P01325 mus musculu
35	41	41.4	195	1	RERB_ARATH	O48671 arabidopsis
36	41	41.4	271	1	YM48_MYCTU	Q10526 mycobacteri
37	41	41.4	306	1	PYRD_METJA	Q58070 methanococc
38	41	41.4	449	1	PCO1_HUMAN	Q15113 homo sapien
39	41	41.4	485	1	MKR4_HUMAN	Q13434 homo sapien
40	41	41.4	759	1	CIQ2_MOUSE	Q9z351 mus musculu
41	41	41.4	914	1	WA22_MYCTU	O06794 mycobacteri
42	41	41.4	1004	1	ATU2_YEAST	P38995 saccharomyc
43	40	40.4	110	1	INS1_RAT	P01322 rattus norv
44	40	40.4	130	1	RS9_LACLA	Q9cdg7 lactococcus
45	40	40.4	138	1	RS9_ANASP	Q8ypk7 anabaena sp

ALIGNMENTS

RESULT 1
POLG_FMDVA
ID POLG_FMDVA STANDARD; PRT; 2332 AA.
AC P03308; P03312; Q65038; Q65039; Q65040; Q65041; Q65042; Q65043;
AC Q65044; Q65045; Q65046; Q65047;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat
DE proteins VP1 TO VP4; Core proteins X, P14, P41, P19; Genome-linked
DE proteins VPG1 TO VPG3; Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48)].
OS Foot-and-mouth disease virus (strain A12) (Aphthovirus A) (FMDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Aphthovirus.
OX NCBI_TaxID=12114;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85211015; PubMed=2987518;
RA Robertson B.H., Grubman M.J., Weddell G.N., Moore D.M., Welsh J.D.,
RA Fischer T., Dowbenko D.J., Yansura D.G., Small B., Kleid D.G.;
RT "Nucleotide and amino acid sequence coding for polypeptides of
RT foot-and-mouth disease virus type A12.";
RL J. Virol. 54:651-660(1985).
RN [2]
RP SEQUENCE OF 1863-2332 FROM N.A.
RX MEDLINE=83225613; PubMed=6305004;
RA Robertson B.H., Morgan D.O., Moore D.M., Grubman M.J., Card J.,
RA Fischer T., Weddell G.N., Dowbenko D.J., Yansura D.G.;
RT "Identification of amino acid and nucleotide sequence of the
RT foot-and-mouth disease virus RNA polymerase.";
RL Virology 126:614-623(1983).
RN [3]
RP SEQUENCE OF 715-955 FROM N.A.
RX MEDLINE=82061853; PubMed=6272395;
RA Kleid D.G., Yansura D.G., Small B., Dowbenko D.J., Moore D.M.,
RA Grubman M.J., McKercher P.D., Morgan D.O., Robertson B.H.,
RA Bachrach H.L.;
RT "Cloned viral protein vaccine for foot-and-mouth disease: responses
RT in cattle and swine.";
RL Science 214:1125-1129(1981).
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC ---

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M10975; AAA42593.1; -.
DR   EMBL; J02187; AAA42670.1; -.
DR   MEROPS; C03.008; -.
DR   InterPro; IPR004004; Calici_pol_hel.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   InterPro; IPR000605; RNA_helicase.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR001205; RNA_pol_P3D.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00073; rhv; 3.
DR   Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR   Pfam; PF00910; RNA_helicase; 1.
DR   PRINTS; PR00918; CALICVIRUSNS.
DR   PRINTS; PR01542; FMDVP1COAT.
KW   Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;
KW   Transferase; Hydrolase; Thiol protease; Nonstructural protein;
KW   Myristate.
FT   CHAIN         1    200       NONSTRUCTURAL PROTEIN P20A.
FT   CHAIN       201    285       COAT PROTEIN VP4.
FT   CHAIN       286    503       COAT PROTEIN VP2.
FT   CHAIN       504    723       COAT PROTEIN VP3.
FT   CHAIN       724    937       COAT PROTEIN VP1.
FT   CHAIN       938    953       CORE PROTEIN X.
FT   CHAIN       954   1107       CORE PROTEIN P14.
FT   CHAIN      1108   1425       CORE PROTEIN P41.
FT   CHAIN      1426   1578       CORE PROTEIN P19.
FT   CHAIN      1579   1601       GENOME-LINKED PROTEIN VPG1.
FT   CHAIN      1602   1625       GENOME-LINKED PROTEIN VPG2.
FT   CHAIN      1626   1649       GENOME-LINKED PROTEIN VPG3.
FT   CHAIN      1650   1862       PROTEASE.
FT   CHAIN      1863   2332       RNA-DIRECTED RNA POLYMERASE.
FT   LIPID       201    201       MYRISTATE.
SQ   SEQUENCE   2332 AA;  259408 MW;  EE77DA739CBEDC6A CRC64;

  Query Match             77.8%;  Score 77;  DB 1;  Length 2332;
  Best Local Similarity   89.5%;  Pred. No. 0.002;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||||| |||| ||||||
Db        864 GSGVRGDFGSLAPRVARQL 882


RESULT 2
POLG_FMDVC
ID   POLG_FMDVC     STANDARD;      PRT;   216 AA.
AC   P03309;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Genome polyprotein [Contains: Coat protein VP1; Core protein P52]
DE   (Fragment).
OS   Foot-and-mouth disease virus (strain A24 Cruzeiro) (Aphthovirus A)
OS   (FMDV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12115;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83143292; PubMed=6298715;
RA   Makoff A.J., Paynter C.A., Rowlands D.J., Boothroyd J.C.;
RT   "Comparison of the amino acid sequence of the major immunogen from
RT   three serotypes of foot and mouth disease virus.";
RL   Nucleic Acids Res. 10:8285-8295(1982).
CC   -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC       EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC       VP3, AND VP4.
CC   -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; J02183; AAA42596.1; -.
DR   EMBL; A06733; CAA00589.1; -.
DR   PIR; A03911; A03911.
DR   HSSP; Q88571; 1TME.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
KW   Coat protein; Core protein; Polyprotein.
FT   NON_TER       1      1
FT   CHAIN         1    202       COAT PROTEIN VP1.
FT   CHAIN       203   >216       CORE PROTEIN P52.
FT   NON_TER     216    216
SQ   SEQUENCE   216 AA;  23889 MW;  501659FF031A1D85 CRC64;

  Query Match             66.7%;  Score 66;  DB 1;  Length 216;
  Best Local Similarity   73.7%;  Pred. No. 0.008;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||| ||| |||| || :||
Db        129 GSGRRGDMGSLAARVVKQL 147


RESULT 3
RS9_STRCO
ID   RS9_STRCO      STANDARD;      PRT;   170 AA.
AC   Q53875;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   30S ribosomal protein S9.
GN   RPSI OR SCO4735 OR SC6G4.13.
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=98102746; PubMed=9439573;
RA   Sanchez C., Blanco G., Mendez C., Salas J.A.;
RT   "Cloning, sequencing and transcriptional analysis of a Streptomyces
RT   coelicolor operon containing the rplM and rpsI genes encoding
RT   ribosomal proteins ScoL13 and ScoS9.";
RL   Mol. Gen. Genet. 257:91-96(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2) / M145;
RX   MEDLINE=21996410; PubMed=12000953;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
CC   -!- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.
CC   ----------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U43429; AAC46061.1; -.
DR   EMBL; AL939121; CAA20391.1; -.
DR   PIR; T35564; T35564.
DR   HAMAP; MF_00532; -; 1.
DR   InterPro; IPR000754; Ribosomal_S9.
DR   Pfam; PF00380; Ribosomal_S9; 1.
DR   ProDom; PD001627; Ribosomal_S9; 1.
DR   PROSITE; PS00360; RIBOSOMAL_S9; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   170 AA;  18699 MW;  3CEE00FF69151C99 CRC64;

  Query Match             50.5%;  Score 50;  DB 1;  Length 170;
  Best Local Similarity   57.9%;  Pred. No. 1.5;
  Matches   11;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              | || | :|:| | ||| |
Db        108 GGGVSGQAGALRLGVARAL 126


RESULT 4
RS9_CHLTE
ID   RS9_CHLTE      STANDARD;      PRT;   129 AA.
AC   Q8KBK5;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   30S ribosomal protein S9.
GN   RPSI OR CT1782.
OS   Chlorobium tepidum.
OC   Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC   Chlorobium.
OX   NCBI_TaxID=1097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TLS / ATCC 49652 / DSM 12025;
RX   MEDLINE=22103685; PubMed=12093901;
RA   Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA   Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
RA   Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
RA   Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA   Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
RA   Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA   Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT   "The complete genome sequence of Chlorobium tepidum TLS, a
RT   photosynthetic, anaerobic, green-sulfur bacterium.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
CC   -!- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AE012931; AAM73003.1; -.
DR   TIGR; CT1782; -.
DR   HAMAP; MF_00532; -; 1.
DR   InterPro; IPR000754; Ribosomal_S9.
DR   Pfam; PF00380; Ribosomal_S9; 1.
DR   ProDom; PD001627; Ribosomal_S9; 1.
DR   PROSITE; PS00360; RIBOSOMAL_S9; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   129 AA;  14541 MW;  D110858BA87B9BAC CRC64;

  Query Match             49.5%;  Score 49;  DB 1;  Length 129;
  Best Local Similarity   47.4%;  Pred. No. 1.6;
  Matches    9;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              | |: | ||:::| :|| |
Db         68 GGGLTGQSGAVSLAIARAL 86


RESULT 5
POLG_FMDV5
ID   POLG_FMDV5     STANDARD;      PRT;   230 AA.
AC   P03307;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Genome polyprotein [Contains: Coat proteins VP3, VP1; Core protein
DE   P52] (Fragment).
OS   Foot-and-mouth disease virus (strain A5) (Aphthovirus A) (FMDV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12113;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84028562; PubMed=6194987;
RA   Beck E., Feil G., Strohmaier K.;
RT   "The molecular basis of the antigenic variation of foot-and-mouth
RT   disease virus.";
RL   EMBO J. 2:555-559(1983).
CC   -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC       EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC       VP3, AND VP4.
CC   -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; V01135; CAA24365.1; ALT_INIT.
DR   EMBL; V01135; CAA24366.1; ALT_SEQ.
DR   PIR; A03909; A03909.
DR   HSSP; Q88571; 1TME.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
KW   Coat protein; Core protein; Polyprotein.
FT   NON_TER       1      1
FT   CHAIN        <1      4       COAT PROTEIN VP3.
FT   CHAIN         5    216       COAT PROTEIN VP1.
FT   CHAIN       217   >230       CORE PROTEIN P52.
FT   NON_TER     230    230
SQ   SEQUENCE   230 AA;  25369 MW;  EA93A190F4CC1608 CRC64;

  Query Match             49.5%;  Score 49;  DB 1;  Length 230;
  Best Local Similarity   57.9%;  Pred. No. 2.9;
  Matches   11;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              |   ||| || | | |:||
Db        143 GGPRRGDMGSAAARAAKQL 161


RESULT 6
POLG_FMDVZ
ID   POLG_FMDVZ     STANDARD;      PRT;  2336 AA.
AC   P49303;
```

```
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Genome polyprotein [Contains: Nonstructural protein P20A; Coat
DE   proteins VP1 TO VP4; Core proteins X, P14, P41, P19; Genome-linked
DE   proteins VPG1 TO VPG3; Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE   (P3C); RNA-directed RNA polymerase (EC 2.7.7.48)].
OS   Foot-and-mouth disease virus (strain A22/550 Azerbaijan 65)
OS   (Aphthovirus A) (FMDV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=73481;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sosnovtsev S.V., Onischenko A.M., Petrov N.A., Kalashnikova T.I.,
RA   Mamaeva N.V., Drygin V.Y., Perevozchikova N.A., Vasilenko S.K.;
RL   Submitted (AUG-1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the
CC       poliovirus polyprotein. In other picornavirus reactions Glu may be
CC       substituted for Gln, and Ser or Thr for Gly.
CC   -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC       {RNA}(N).
CC   -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC       EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC       VP3, AND VP4.
CC   -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC   -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X74812; CAA52812.1; -.
DR   PIR; S37077; S37077.
DR   HSSP; Q88571; 1TME.
DR   MEROPS; C03.008; -.
DR   MEROPS; C28.001; -.
DR   InterPro; IPR004004; Calici_pol_hel.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   InterPro; IPR000605; RNA_helicase.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR001205; RNA_pol_P3D.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00073; rhv; 3.
DR   Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR   Pfam; PF00910; RNA_helicase; 1.
DR   PRINTS; PR00918; CALICVIRUSNS.
DR   PRINTS; PR01542; FMDVP1COAT.
KW   Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;
KW   Transferase; Hydrolase; Thiol protease; Nonstructural protein;
KW   Myristate.
FT   CHAIN        1    201       NONSTRUCTURAL PROTEIN P20A.
FT   CHAIN      202    286       COAT PROTEIN VP4.
FT   CHAIN      287    504       COAT PROTEIN VP2.
FT   CHAIN      505    724       COAT PROTEIN VP3.
FT   CHAIN      725    938       COAT PROTEIN VP1.
FT   CHAIN      939    954       CORE PROTEIN X.
FT   CHAIN      955   1108       CORE PROTEIN P14.
FT   CHAIN     1109   1426       CORE PROTEIN P41.
FT   CHAIN     1427   1579       CORE PROTEIN P19.
FT   CHAIN     1580   1602       GENOME-LINKED PROTEIN VPG1.
FT   CHAIN     1603   1626       GENOME-LINKED PROTEIN VPG2.
FT   CHAIN     1627   1650       GENOME-LINKED PROTEIN VPG3.
FT   CHAIN     1651   1863       PROTEASE.
FT   CHAIN     1864   2336       RNA-DIRECTED RNA POLYMERASE.
FT   LIPID      202    202       MYRISTATE (BY SIMILARITY).
SQ   SEQUENCE   2336 AA;  259982 MW;  15AC2AB022B5B954 CRC64;
```

```
  Query Match            48.5%;  Score 48;  DB 1;  Length 2336;
  Best Local Similarity  63.2%;  Pred. No. 42;
  Matches   12;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              | | |||   || ||| ||
Db        865 GMGRRGDLEPLAARVAAQL 883


RESULT 7
RS9_THEMA
ID   RS9_THEMA      STANDARD;      PRT;   134 AA.
AC   Q9X1G4;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   30S ribosomal protein S9.
GN   RPSI OR TM1453.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=99287316; PubMed=10360571;
RA   Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA   Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA   McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA   Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA   Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA   Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT   "Evidence for lateral gene transfer between Archaea and Bacteria from
RT   genome sequence of Thermotoga maritima.";
RL   Nature 399:323-329(1999).
CC   -!- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001797; AAD36521.1; -.
DR   PIR; F72250; F72250.
DR   TIGR; TM1453; -.
DR   HAMAP; MF_00532; -; 1.
DR   InterPro; IPR000754; Ribosomal_S9.
DR   Pfam; PF00380; Ribosomal_S9; 1.
DR   ProDom; PD001627; Ribosomal_S9; 1.
DR   PROSITE; PS00360; RIBOSOMAL_S9; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   134 AA;  15198 MW;  698FB55B73E0A539 CRC64;

  Query Match            46.5%;  Score 46;  DB 1;  Length 134;
  Best Local Similarity  47.4%;  Pred. No. 4.7;
  Matches    9;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              | |: | ||:: | :|| |
Db         73 GGGLSGQSGAVRLGIARAL 91


RESULT 8
TRAD_HUMAN
ID   TRAD_HUMAN     STANDARD;      PRT;   312 AA.
AC   Q15628;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Tumor necrosis factor receptor type 1 associated DEATH domain protein
DE   (TNFR1-associated DEATH domain protein) (TNFRSF1A-associated via death
```

```
DE   domain).
GN   TRADD.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95277836; PubMed=7758105;
RA   Hsu H., Xiong J., Goeddel D.V.;
RT   "The TNF receptor 1-associated protein TRADD signals cell death and
RT   NF-kappa B activation.";
RL   Cell 81:495-504(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Scheuerpflug C.G., Dechant M., Fellenberg J., Ewerbeck V.,
RA   Debatin K.M.;
RT   "Sequence, genomic organisation, and mutation analysis of the human
RT   TRADD gene in childhood B- and T-lineage acute lymphoblastic leukemia
RT   and ALPS.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreas;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   INTERACTION WITH RIPK1.
RX   MEDLINE=96200892; PubMed=8612133;
RA   Hsu H., Huang J., Shu H.-B., Baichwal V.R., Goeddel D.V.;
RT   "TNF-dependent recruitment of the protein kinase RIP to the TNF
RT   receptor-1 signaling complex.";
RL   Immunity 4:387-396(1996).
RN   [5]
RP   STRUCTURE BY NMR OF 1-179.
RX   PubMed=10911999;
RA   Tsao D.H., McDonagh T., Telliez J.B., Hsu S., Malakian K., Xu G.Y.,
RA   Lin L.L.;
RT   "Solution structure of N-TRADD and characterization of the interaction
RT   of N-TRADD and C-TRAF2, a key step in the TNFR1 signaling pathway.";
RL   Mol. Cell 5:1051-1057(2000).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 1-179 IN COMPLEX WITH
RP   TRAF2.
RX   PubMed=10892748;
RA   Park Y.C., Ye H., Hsia C., Segal D., Rich R.L., Liou H.C.,
RA   Myszka D.G., Wu H.;
RT   "A novel mechanism of TRAF signaling revealed by structural and
RT   functional analyses of the TRADD-TRAF2 interaction.";
RL   Cell 101:777-787(2000).
CC   -!- FUNCTION: Adapter molecule for TNFRSF1A/TNFR1 that specifically
CC       associates with the cytoplasmic domain of activated TNFRSF1A/TNFR1
CC       mediating its interaction with FADD. Overexpression of TRADD leads
CC       to two major TNF-induced responses, apoptosis and activation of
CC       NF-kappa-B.
CC   -!- SUBUNIT: Heterodimer with TNFRSF1A/TNFR1. Binds to TRAF2.
CC       Interacts with TRAF1, FADD and RIPK1.
CC   -!- TISSUE SPECIFICITY: Found in all examined tissues.
CC   -!- DOMAIN: Requires the intact DEATH domain to associate with
CC       TNFRSF1A/TNFR1.
CC   -!- SIMILARITY: Contains 1 death domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L41690; AAA98482.1; ALT_INIT.
DR   EMBL; AJ311614; CAC38018.2; -.
DR   EMBL; AJ311615; CAC38018.2; JOINED.
DR   EMBL; AJ311616; CAC38018.2; JOINED.
DR   EMBL; BC004491; AAH04491.1; -.
DR   PIR; A56911; A56911.
DR   PDB; 1F2H; 24-MAY-00.
DR   PDB; 1F3V; 06-JUN-00.
DR   Genew; HGNC:12030; TRADD.
DR   MIM; 603500; -.
DR   GO; GO:0006917; P:induction of apoptosis; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR000488; Death.
DR   Pfam; PF00531; death; 1.
DR   SMART; SM00005; DEATH; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Apoptosis; 3D-structure.
FT   DOMAIN      179    289       DEATH.
FT   DOMAIN      192    198       POLY-PRO.
SQ   SEQUENCE   312 AA;  34247 MW;  5645D7E63E5FF05A CRC64;

  Query Match             46.5%;  Score 46;  DB 1;  Length 312;
  Best Local Similarity   61.5%;  Pred. No. 11;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDSGSLA 13
              |||| ||  | :|
Db        164 CGSGARGGDGEVA 176


RESULT 9
POLG_FMDV1
ID   POLG_FMDV1     STANDARD;      PRT;  2333 AA.
AC   P03306; Q64768; Q84750; Q84751; Q84752; Q84753; Q84754; Q84760;
AC   Q84761; Q84762; Q84763; Q84764; Q84765; Q84766; Q84767; Q84768;
AC   Q84769; Q89824;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Genome polyprotein [Contains: Nonstructural protein P20A; Coat
DE   proteins VP1 TO VP4; Core protein P52; Genome-linked proteins VPG1 TO
DE   VPG3; Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed
DE   RNA polymerase P56A (EC 2.7.7.48)].
OS   Foot-and-mouth disease virus (strain A10-61) (Aphthovirus A) (FMDV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12112;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84169547; PubMed=6324120;
RA   Carroll A.R., Rowlands D.J., Clarke B.E.;
RT   "The complete nucleotide sequence of the RNA coding for the primary
RT   translation product of foot and mouth disease virus.";
RL   Nucleic Acids Res. 12:2461-2472(1984).
RN   [2]
RP   SEQUENCE OF 115-1048 FROM N.A.
RX   MEDLINE=82211814; PubMed=6282711;
```

```
RA   Boothroyd J.C., Harris T.J.R., Rowlands D.J., Lowe P.A.;
RT   "The nucleotide sequence of cDNA coding for the structural proteins
RT   of foot-and-mouth disease virus.";
RL   Gene 17:153-161(1982).
CC   -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the
CC       poliovirus polyprotein. In other picornavirus reactions Glu may be
CC       substituted for Gln, and Ser or Thr for Gly.
CC   -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC       {RNA}(N).
CC   -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC       EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC       VP3, AND VP4.
CC   -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC   -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V01130; CAA24361.1; -.
DR   EMBL; X00429; CAA25127.1; -.
DR   MEROPS; C03.008; -.
DR   MEROPS; C28.001; -.
DR   InterPro; IPR004004; Calici_pol_hel.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   InterPro; IPR000605; RNA_helicase.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR001205; RNA_pol_P3D.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00073; rhv; 3.
DR   Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR   Pfam; PF00910; RNA_helicase; 1.
DR   PRINTS; PR00918; CALICVIRUSNS.
DR   PRINTS; PR01542; FMDVP1COAT.
KW   Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;
KW   Transferase; Hydrolase; Thiol protease; Nonstructural protein;
KW   Myristate.
FT   CHAIN         1    201       NONSTRUCTURAL PROTEIN P20A.
FT   CHAIN       202    286       COAT PROTEIN VP4.
FT   CHAIN       287    504       COAT PROTEIN VP2.
FT   CHAIN       505    725       COAT PROTEIN VP3.
FT   CHAIN       726    937       COAT PROTEIN VP1.
FT   CHAIN       938   1578       CORE PROTEIN P52.
FT   CHAIN      1579   1601       GENOME-LINKED PROTEIN VPG1.
FT   CHAIN      1602   1625       GENOME-LINKED PROTEIN VPG2.
FT   CHAIN      1626   1649       GENOME-LINKED PROTEIN VPG3.
FT   CHAIN      1650   1863       PROTEASE P20B.
FT   CHAIN      1864   2333       RNA-DIRECTED RNA POLYMERASE P56A.
FT   LIPID       202    202       MYRISTATE.
FT   CONFLICT    396    396       S -> C (IN REF. 2).
FT   CONFLICT    632    632       P -> L (IN REF. 2).
SQ   SEQUENCE   2333 AA;  259645 MW;  4FC667DCC521BC60 CRC64;

  Query Match             46.5%;  Score 46;  DB 1;  Length 2333;
  Best Local Similarity   71.4%;  Pred. No. 83;
  Matches   10;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 GDSGSLALRVARQL 20
              || ||:| ||| ||
Db        869 GDLGSIAARVATQL 882


RESULT 10
AMPA_BUCBP
ID   AMPA_BUCBP     STANDARD;      PRT;   501 AA.
AC   Q89AG2;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Probable cytosol aminopeptidase (EC 3.4.11.1) (Leucine aminopeptidase)
DE   (LAP) (Leucyl aminopeptidase).
GN   PEPA OR BBP332.
OS   Buchnera aphidicola (subsp. Baizongia pistaciae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Buchnera.
OX   NCBI_TaxID=135842;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426901; PubMed=12522265;
RA   Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,
RA   Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,
RA   Tamames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.;
RT   "Reductive genome evolution in Buchnera aphidicola.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).
CC   -!- FUNCTION: Presumably involved in the processing and regular
CC       turnover of intracellular proteins. Catalyzes the removal of
CC       unsubstituted N-terminal amino acids from various peptides (By
CC       similarity).
CC   -!- CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa-|-
CC       Xbb-, in which Xaa is preferably Leu, but may be other amino acids
CC       including Pro although not Arg or Lys, and Xbb may be Pro.
CC   -!- COFACTOR: Binds 2 manganese ions per subunit (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M17.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE014017; AAO27053.1; -.
DR   HAMAP; MF_00181; -; 1.
DR   PROSITE; PS00631; CYTOSOL_AP; 1.
KW   Hydrolase; Aminopeptidase; Manganese; Complete proteome.
FT   ACT_SITE    284    284       POTENTIAL.
FT   ACT_SITE    358    358       POTENTIAL.
FT   METAL       272    272       MANGANESE 2 (BY SIMILARITY).
FT   METAL       277    277       MANGANESE 1 AND 2 (BY SIMILARITY).
FT   METAL       295    295       MANGANESE 2 (BY SIMILARITY).
FT   METAL       354    354       MANGANESE 1 (BY SIMILARITY).
FT   METAL       356    356       MANGANESE 1 AND 2 (BY SIMILARITY).
SQ   SEQUENCE   501 AA;  56088 MW;  F0628DD186BCB80E CRC64;

  Query Match             45.5%;  Score 45;  DB 1;  Length 501;
  Best Local Similarity   42.1%;  Pred. No. 25;
  Matches    8;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              | ||  ||| :::: :| |
Db        271 GKGVTFDSGGISIKTSRDL 289


RESULT 11
AMY1_DEBOC
ID   AMY1_DEBOC     STANDARD;      PRT;   512 AA.
AC   P19269;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Alpha-amylase 1 precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE   glucanohydrolase).
GN   AMY1.
OS   Debaryomyces occidentalis (Yeast) (Schwanniomyces occidentalis).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX   NCBI_TaxID=27300;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=ATCC 26076;
RX   MEDLINE=90032659; PubMed=2806251;
RA   Strasser A.W.M., Selk R., Dohmen R.J., Niermann T., Bielefeld M.,
RA   Seeboth P., Tu G., Hollenberg C.P.;
RT   "Analysis of the alpha-amylase gene of Schwanniomyces occidentalis
RT   and the secretion of its gene product in transformants of different
RT   yeast genera.";
RL   Eur. J. Biochem. 184:699-706(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CCRC 21164;
RX   MEDLINE=92120467; PubMed=1769525;
RA   Wu F.M., Wang T.T., Hsu W.H.;
RT   "The nucleotide sequence of Schwanniomyces occidentalis alpha-amylase
RT   gene.";
RL   FEMS Microbiol. Lett. 66:313-318(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 26077 / CBS 2863;
RX   MEDLINE=92307400; PubMed=1612414;
RA   Park J.C., Bai S., Tai C.Y., Chun S.B.;
RT   "Nucleotide sequence of the extracellular alpha-amylase gene in the
RT   yeast Schwanniomyces occidentalis ATCC 26077.";
RL   FEMS Microbiol. Lett. 72:17-23(1992).
CC   -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC       linkages in oligosaccharides and polysaccharides.
CC   -!- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC   -!- ENZYME REGULATION: ALPHA-AMYLASE EXPRESSION UNDERLIES CATABOLITE
CC       REPRESSION BY GLUCOSE.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S77586; AAB21151.2; -.
DR   EMBL; X16040; CAA34162.1; -.
DR   EMBL; X62079; CAA43995.1; -.
DR   EMBL; S38381; AAB22383.2; -.
DR   PIR; S06115; S06115.
DR   PIR; S23355; S23355.
DR   HSSP; P10529; 7TAA.
DR   InterPro; IPR006589; Alp_amyl_cat_sub.
DR   InterPro; IPR006047; Alpha_amyl_cat.
DR   InterPro; IPR006046; Glyco_hydro_13.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   PRINTS; PR00110; ALPHAAMYLASE.
DR   SMART; SM00642; Aamy; 1.
KW   Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
KW   Glycoprotein; Signal.
FT   SIGNAL       1     25       POTENTIAL.
FT   CHAIN       26    512       ALPHA-AMYLASE 1.
FT   ACT_SITE   242    242       BY SIMILARITY.
FT   ACT_SITE   246    246       BY SIMILARITY.
FT   ACT_SITE   333    333       BY SIMILARITY.
FT   CARBOHYD   233    233       N-LINKED (GLCNAC. . .) (PROBABLE).
FT   DISULFID    66     74       BY SIMILARITY.
FT   DISULFID   186    200       BY SIMILARITY.
FT   DISULFID   276    319       BY SIMILARITY.
FT   DISULFID   475    510       BY SIMILARITY.
FT   VARIANT     32     32       M -> K (IN STRAINS CCRC 21164 AND ATCC
FT                               26077).
FT   VARIANT     36     36       S -> G (IN STRAIN CCRC 21164).
FT   VARIANT     73     73       Y -> I (IN STRAIN ATCC 26077).
FT   VARIANT    280    280       N -> S (IN STRAIN CCRC 21164).
FT   VARIANT    350    350       D -> A (IN STRAINS CCRC 21164 AND ATCC
FT                               26077).
FT   VARIANT    479    479       L -> S (IN STRAINS CCRC 21164 AND ATCC
FT                               26077).
FT   VARIANT    483    483       S -> F (IN STRAIN CCRC 21164).
SQ   SEQUENCE   512 AA;  56527 MW;  857552B2CF60F965 CRC64;

  Query Match             45.5%;  Score 45;  DB 1;  Length 512;
  Best Local Similarity   47.1%;  Pred. No. 25;
  Matches    8;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDSGSLALRVA 17
              | | : ||||||:: ::
Db        475 CNSVLAGDSGSLSVSIS 491


RESULT 12
CH11_RHOSH
ID   CH11_RHOSH     STANDARD;      PRT;    94 AA.
AC   P25969; Q59772;
DT   01-MAY-1992 (Rel. 22, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   10 kDa chaperonin 1 (Protein Cpn10 1) (groES protein 1).
GN   GROS1 OR GROES1 OR GROES.
OS   Rhodobacter sphaeroides (Rhodopseudomonas sphaeroides).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC   Rhodobacteraceae; Rhodobacter.
OX   NCBI_TaxID=1063;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HR;
RX   MEDLINE=97144535; PubMed=8990302;
RA   Lee W.T., Terlesky K.C., Tabita F.R.;
RT   "Cloning and characterization of two groESL operons of Rhodobacter
RT   sphaeroides: transcriptional regulation of the heat-induced groESL
RT   operon.";
RL   J. Bacteriol. 179:487-495(1997).
RN   [2]
RP   SEQUENCE OF 1-22.
RX   MEDLINE=91329399; PubMed=1678280;
RA   Terlesky K.C., Tabita F.R.;
RT   "Purification and characterization of the chaperonin 10 and
RT   chaperonin 60 proteins from Rhodobacter sphaeroides.";
RL   Biochemistry 30:8181-8186(1991).
CC   -!- FUNCTION: Binds to Cpn60 in the presence of Mg-ATP and suppresses
CC       the ATPase activity of the latter.
CC   -!- SUBUNIT: Heptamer of 7 subunits arranged in a ring (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- INDUCTION: By heat shock.
CC   -!- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U37369; AAB41335.1; -.
DR   HAMAP; MF_00580; -; 1.
DR   InterPro; IPR001476; Chaprnin_Cpn10.
DR   Pfam; PF00166; cpn10; 1.
DR   PRINTS; PR00297; CHAPERONIN10.
DR   ProDom; PD000566; Chaprnin_Cpn10; 1.
DR   PROSITE; PS00681; CHAPERONINS_CPN10; 1.
KW   Chaperone; Heat shock.
FT   INIT_MET     0      0
SQ   SEQUENCE   94 AA;  10065 MW;  98B712311DABA8C6 CRC64;

  Query Match             44.4%;  Score 44;  DB 1;  Length 94;
  Best Local Similarity   66.7%;  Pred. No. 6.5;
  Matches    8;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy          1 CGSGVRGDSGSL 12
              || | | ||| |
Db         42 CGEGARKDSGEL 53


RESULT 13
CH10_RHOPA
ID   CH10_RHOPA     STANDARD;      PRT;    95 AA.
AC   Q93MH2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   10 kDa chaperonin (Protein Cpn10) (groES protein).
GN   GROS OR GROES.
OS   Rhodopseudomonas palustris.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Bradyrhizobiaceae; Rhodopseudomonas.
OX   NCBI_TaxID=1076;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Xiao M., Zhu C.R., Qian X.M., Zheng P., Chen Y.Y.;
RT   "Cloning and sequencing of the groESL Operon of Rhodopseudomonas
RT   palustris.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds to Cpn60 in the presence of Mg-ATP and suppresses
CC       the ATPase activity of the latter.
CC   -!- SUBUNIT: Heptamer of 7 subunits arranged in a ring (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF406639; AAK94942.1; -.
DR   HAMAP; MF_00580; -; 1.
DR   InterPro; IPR001476; Chaprnin_Cpn10.
DR   Pfam; PF00166; cpn10; 1.
DR   PRINTS; PR00297; CHAPERONIN10.
DR   ProDom; PD000566; Chaprnin_Cpn10; 1.
DR   PROSITE; PS00681; CHAPERONINS_CPN10; 1.
KW   Chaperone.
SQ   SEQUENCE   95 AA;  10181 MW;  EA3C1BA1E444015B CRC64;

  Query Match             44.4%;  Score 44;  DB 1;  Length 95;
  Best Local Similarity   66.7%;  Pred. No. 6.6;
  Matches    8;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDSGSL 12
              || | | ||| |
Db         43 CGEGARKDSGEL 54


RESULT 14
RR9_CYACA
ID   RR9_CYACA      STANDARD;      PRT;   128 AA.
AC   Q9TLV4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Chloroplast 30S ribosomal protein S9.
GN   RPS9.
OS   Cyanidium caldarium.
OG   Chloroplast.
OC   Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
OC   Cyanidium.
OX   NCBI_TaxID=2771;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RK-1;
RX   MEDLINE=20496959; PubMed=11040290;
RA   Gloeckner G., Rosenthal A., Valentin K.-U.;
RT   "The structure and gene repertoire of an ancient red algal plastid
RT   genome.";
RL   J. Mol. Evol. 51:382-390(2000).
CC   -!- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF022186; AAF12930.1; -.
DR   HAMAP; MF_00532; -; 1.
DR   InterPro; IPR000754; Ribosomal_S9.
DR   Pfam; PF00380; Ribosomal_S9; 1.
DR   ProDom; PD001627; Ribosomal_S9; 1.
DR   PROSITE; PS00360; RIBOSOMAL_S9; FALSE_NEG.
KW   Ribosomal protein; Chloroplast.
SQ   SEQUENCE   128 AA;  14329 MW;  DFFAA6345D338C8B CRC64;

  Query Match             44.4%;  Score 44;  DB 1;  Length 128;
  Best Local Similarity   42.1%;  Pred. No. 8.9;
  Matches    8;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              | |: | :|:: | :|| |
Db         67 GGGLSGQAGAIRLAIARVL 85


RESULT 15
RS9_MYCPU
ID   RS9_MYCPU      STANDARD;      PRT;   136 AA.
AC   Q98Q72;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   30S ribosomal protein S9.
GN   RPSI OR MYPU_4960.
OS   Mycoplasma pulmonis.
OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2107;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=UAB CTIP;
RX   MEDLINE=21267165; PubMed=11353084;
RA   Chambaud I., Heilig R., Ferris S., Barbe V., Samson D.; Galisson F.,
RA   Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,
RA   Blanchard A.;
RT   "The complete genome sequence of the murine respiratory pathogen
RT   Mycoplasma pulmonis.";
RL   Nucleic Acids Res. 29:2145-2153(2001)..
CC   -!- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL445564; CAC13669.1; -.
DR   PIR; H90573; H90573.
DR   MypuList; MYPU_4960; -.
DR   HAMAP; MF_00532; -; 1.
DR   InterPro; IPR000754; Ribosomal_S9.
DR   Pfam; PF00380; Ribosomal_S9; 1.
```

```
DR   ProDom; PD001627; Ribosomal_S9; 1.
DR   PROSITE; PS00360; RIBOSOMAL_S9; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   136 AA;  15268 MW;  78EBE9CC47D51DF1 CRC64;

  Query Match           44.4%;  Score 44;  DB 1;  Length 136;
  Best Local Similarity  42.1%;  Pred. No. 9.4;
  Matches    8;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              | |: | :|:: | :|| |
Db         75 GGGLAGQAGAIRLGIARAL 93


Search completed: October  9, 2003, 10:15:00
Job time : 12.3333 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:09:40 ; Search time 49.3333 Seconds
(without alignments)
104.616 Million cell updates/sec

Title: US-09-549-186B-9
Perfect score: 99
Sequence: 1 CGSGVRGDSGSLALRVARQL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	83	83.8	20	12	Q9PXE4	Q9pxe4 foot-and-mo
2	70	70.7	213	12	Q912L2	Q91212 foot-and-mo
3	70	70.7	213	12	Q67446	Q67446 foot-and-mo
4	70	70.7	213	12	Q67448	Q67448 foot-and-mo
5	70	70.7	213	12	Q65093	Q65093 foot-and-mo
6	67	67.7	213	12	Q98VZ6	Q98vz6 foot-and-mo
7	66	66.7	210	12	Q67438	Q67438 foot-and-mo
8	66	66.7	969	12	Q9Q2N6	Q9q2n6 foot-and-mo
9	64	64.6	126	12	Q8V443	Q8v443 foot-and-mo
10	64	64.6	169	12	Q67442	Q67442 foot-and-mo
11	60	60.6	125	12	Q8V442	Q8v442 foot-and-mo
12	60	60.6	143	12	Q9YPU0	Q9ypu0 foot-and-mo
13	60	60.6	208	12	Q8JUR8	Q8jur8 foot-and-mo
14	60	60.6	211	12	Q8JUP0	Q8jup0 foot-and-mo
15	60	60.6	212	12	Q8JUP1	Q8jup1 foot-and-mo
16	60	60.6	213	12	Q8JUS6	Q8jus6 foot-and-mo
17	60	60.6	213	12	Q8JUM1	Q8jum1 foot-and-mo
18	60	60.6	213	12	Q8JUQ9	Q8juq9 foot-and-mo
19	60	60.6	213	12	Q8JUP7	Q8jup7 foot-and-mo
20	60	60.6	213	12	Q8JUP4	Q8jup4 foot-and-mo
21	60	60.6	213	12	Q912L1	Q912l1 foot-and-mo
22	60	60.6	213	12	Q8JUQ5	Q8juq5 foot-and-mo
23	60	60.6	213	12	Q8JUL1	Q8jul1 foot-and-mo
24	60	60.6	213	12	Q8JUN1	Q8jun1 foot-and-mo
25	60	60.6	213	12	Q8JUP5	Q8jup5 foot-and-mo
26	60	60.6	213	12	Q8JUN0	Q8jun0 foot-and-mo
27	60	60.6	213	12	Q8JUP3	Q8jup3 foot-and-mo
28	60	60.6	213	12	Q8JUQ2	Q8juq2 foot-and-mo
29	60	60.6	213	12	Q8JUQ1	Q8juq1 foot-and-mo
30	60	60.6	213	12	Q8JUL0	Q8jul0 foot-and-mo
31	60	60.6	213	12	Q8JUS7	Q8jus7 foot-and-mo
32	60	60.6	213	12	Q8JUM0	Q8jum0 foot-and-mo
33	60	60.6	213	12	Q8JUM3	Q8jum3 foot-and-mo
34	60	60.6	213	12	Q8JUQ8	Q8juq8 foot-and-mo
35	60	60.6	213	12	Q8JUN9	Q8jun9 foot-and-mo
36	60	60.6	213	12	Q8JUR0	Q8jur0 foot-and-mo
37	60	60.6	213	12	Q8JUS5	Q8jus5 foot-and-mo
38	58	58.6	197	12	Q8JUT2	Q8jut2 foot-and-mo
39	58	58.6	213	12	Q8JUQ4	Q8juq4 foot-and-mo
40	58	58.6	652	12	Q9Q2N8	Q9q2n8 foot-and-mo
41	57	57.6	213	12	Q8JUM9	Q8jum9 foot-and-mo
42	57	57.6	213	12	Q912K6	Q912k6 foot-and-mo
43	57	57.6	213	12	Q66924	Q66924 foot-and-mo
44	57	57.6	213	12	Q912K4	Q912k4 foot-and-mo
45	57	57.6	213	12	Q912L0	Q912l0 foot-and-mo

ALIGNMENTS

RESULT 1
Q9PXE4
ID Q9PXE4 PRELIMINARY; PRT; 20 AA.
AC Q9PXE4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE VP1 protein (Fragment).
OS Foot-and-mouth disease virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Aphthovirus.
OX NCBI_TaxID=12110;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96100820; PubMed=7483796;
RA Piatti P., Hassard S., Newman J.F., Brown F.;
RT "Antigenic variants in a plaque-isolate of foot-and-mouth disease
RT virus: implications for vaccine production.";
RL Vaccine 13:781-784(1995).
SQ SEQUENCE 20 AA; 1925 MW; E91F87C99C19D7DB CRC64;

Query Match 83.8%; Score 83; DB 12; Length 20;
Best Local Similarity 94.7%; Pred. No. 3.2e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GSGVRGDSGSLALRVARQL 20
||||||||||||||||| |
Db 1 GSGVRGDSGSLALRVARGL 19

RESULT 2
Q912L2
ID Q912L2 PRELIMINARY; PRT; 213 AA.
AC Q912L2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE VP1 protein (Fragment).
GN 1D.

```
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A Arg/68;
RX   MEDLINE=21580796; PubMed=11724271;
RA   Konig G.A., Blanco C., Knowles N.J., Palma E.L., Maradei E.,
RA   Piccone M.E.;
RT   "Phylogenetic analysis of foot-and-mouth disease viruses isolated in
RT   Argentina.";
RL   Virus Genes 23:175-182(2001).
DR   EMBL; AJ308694; CAC48168.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     213    213
SQ   SEQUENCE   213 AA;  23488 MW;  203EFCB8AB45EECE CRC64;

  Query Match             70.7%;  Score 70;  DB 12;  Length 213;
  Best Local Similarity   78.9%;  Pred. No. 0.0048;
  Matches   15;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||| ||| |||| |||:||
Db        140 GSGRRGDMGSLAARVAKQL 158


RESULT 3
Q67446
ID   Q67446      PRELIMINARY;      PRT;   213 AA.
AC   Q67446; Q67447;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Foot and mouth disease virus (Serotype Aven76), capsid protein VP1
DE   mRNA (Fragment).
OS   Foot-and-mouth disease virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12110;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85190547; PubMed=2986125;
RA   Weddell G.N., Yansura D.G., Dowbenko D.J., Hoatlin M.E., Grubman M.J.,
RA   Moore D.M., Kleid D.G.;
RT   "Sequence variation in the gene for the immunogenic capsid protein VP1
RT   of foot-and-mouth disease virus type A.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:2618-2622(1985).
DR   EMBL; K03344; AAA42600.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     213    213
SQ   SEQUENCE   213 AA;  23311 MW;  C4ED4E0116A0DB8A CRC64;

  Query Match             70.7%;  Score 70;  DB 12;  Length 213;
  Best Local Similarity   78.9%;  Pred. No. 0.0048;
  Matches   15;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||| ||| |||| |||:||
Db        140 GSGRRGDMGSLAARVAKQL 158


RESULT 4
Q67448
ID   Q67448      PRELIMINARY;      PRT;   213 AA.
AC   Q67448; Q67449;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Foot and mouth disease virus (Serotype Aarg79), capsid protein VP1
DE   mRNA (Fragment).
OS   Foot-and-mouth disease virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12110;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85190547; PubMed=2986125;
RA   Weddell G.N., Yansura D.G., Dowbenko D.J., Hoatlin M.E., Grubman M.J.,
RA   Moore D.M., Kleid D.G.;
RT   "Sequence variation in the gene for the immunogenic capsid protein VP1
RT   of foot-and-mouth disease virus type A.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:2618-2622(1985).
DR   EMBL; K03345; AAA42601.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     213    213
SQ   SEQUENCE   213 AA;  23345 MW;  546C7FF78CD45CC0 CRC64;

  Query Match             70.7%;  Score 70;  DB 12;  Length 213;
  Best Local Similarity   78.9%;  Pred. No. 0.0048;
  Matches   15;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||| ||| |||| |||:||
Db        140 GSGRRGDMGSLAARVAKQL 158


RESULT 5
Q65093
ID   Q65093      PRELIMINARY;      PRT;   213 AA.
AC   Q65093;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Foot and mouth disease virus strain A Venceslau VP1 (Fragment).
OS   Foot-and-mouth disease virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12110;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85077620; PubMed=6096217;
RA   Cheung A., Whitehead P., Weiss S., Kupper H.;
RT   "Nucleotide sequence of the VP1 gene of the foot-and-mouth disease
RT   virus strain A Venceslau.";
RL   Gene 30:241-245(1984).
DR   EMBL; M12905; AAA42663.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     213    213
SQ   SEQUENCE   213 AA;  23327 MW;  048088D9571D7073 CRC64;

  Query Match             70.7%;  Score 70;  DB 12;  Length 213;
  Best Local Similarity   78.9%;  Pred. No. 0.0048;
  Matches   15;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||| ||| |||| |||:||
Db        140 GSGRRGDMGSLAARVAKQL 158.
```

```
RESULT 6
Q98VZ6
ID   Q98VZ6      PRELIMINARY;      PRT;   213 AA.
AC   Q98VZ6;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   VP1 (Fragment).
GN   1D.
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A76/Argentina/76;
RX   MEDLINE=21580796; PubMed=11724271;
RA   Konig G.A., Blanco C., Knowles N.J., Palma E.L., Maradei E.,
RA   Piccone M.E.;
RT   "Phylogenetic analysis of foot-and-mouth disease viruses isolated in
RT   Argentina.";
RL   Virus Genes 23:175-182(2001).
DR   EMBL; AJ409219; CAC34727.1; -.
DR   HSSP; Q88571; 1TME.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     213    213
SQ   SEQUENCE   213 AA;  23494 MW;  92CBE4DBC885F3C9 CRC64;

  Query Match             67.7%;  Score 67;  DB 12;  Length 213;
  Best Local Similarity   78.9%;  Pred. No. 0.014;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||  ||| |||| ||||||
Db        140 GSARRGDMGSLAARVARQL 158


RESULT 7
Q67438
ID   Q67438      PRELIMINARY;      PRT;   210 AA.
AC   Q67438; Q67439;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Foot and mouth disease virus (Serotype A24), capsid protein VP1 mRNA
DE   (Fragment).
OS   Foot-and-mouth disease virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12110;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85190547; PubMed=2986125;
RA   Weddell G.N., Yansura D.G., Dowbenko D.J., Hoatlin M.E., Grubman M.J.,
RA   Moore D.M., Kleid D.G.;
RT   "Sequence variation in the gene for the immunogenic capsid protein VP1
RT   of foot-and-mouth disease virus type A.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:2618-2622(1985).
DR   EMBL; K03340; AAA42595.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     210    210
SQ   SEQUENCE   210 AA;  23089 MW;  1C96CA1C5E4F6A60 CRC64;

  Query Match             66.7%;  Score 66;  DB 12;  Length 210;
  Best Local Similarity   73.7%;  Pred. No. 0.02;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||| ||| |||| || :||
Db        137 GSGRRGDMGSLAARVVKQL 155


RESULT 8
Q9Q2N6
ID   Q9Q2N6      PRELIMINARY;      PRT;   969 AA.
AC   Q9Q2N6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Polyprotein (Fragment).
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A24/Cruzeiro/Brazil/55;
RA   Abrams C.C.;
RT   "Nucleotide sequence of foot-and-mouth disease virus
RT   A24/Cruzeiro/Brazil/55 from the poly(C) tract to 2B.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A24/Cruzeiro/Brazil/55;
RA   Knowles N.J., Samuel A.R., Aktas S., Rowe C.A., Abrams C.C.,
RA   Newman J.W.I., King A.M.Q.;
RT   "Phylogenetic comparison of the capsid-coding region of all seven
RT   foot-and-mouth disease virus serotypes.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ251476; CAB62583.1; -.
DR   HSSP; Q88571; 1TME.
DR   MEROPS; C28.001; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 3.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   CHAIN         1   >201       LEADER.
FT   CHAIN       202   >286       1A (VP4).
FT   CHAIN       287   >504       1B (VP2).
FT   CHAIN       505   >725       1C (VP3).
FT   CHAIN       726   >938       1D (VP1).
FT   CHAIN       939   >954       2A.
FT   CHAIN       955   >969       2B.
FT   NON_TER     969    969
SQ   SEQUENCE   969 AA;  107629 MW;  7DD5D908FCFF89FD CRC64;

  Query Match             66.7%;  Score 66;  DB 12;  Length 969;
  Best Local Similarity   73.7%;  Pred. No. 0.11;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||| ||| |||| || :||
Db        865 GSGRRGDMGSLAARVVKQL 883


RESULT 9
Q8V443
ID   Q8V443      PRELIMINARY;      PRT;   126 AA.
AC   Q8V443;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   VP1 (Fragment).
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=KEN/1/76;
RA   van Rensburg H.G., Bastos A.D., Haydon D., Joubert F., Nel L.H.;
RT   "Gene heterogeneity in the leader and 3C proteinases of foot-and-mouth
RT   disease virus.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF390861; AAL73359.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     126    126
SQ   SEQUENCE   126 AA;  13693 MW;  82675FAC45F713BB CRC64;

  Query Match             64.6%;  Score 64;  DB 12;  Length 126;
  Best Local Similarity   77.8%;  Pred. No. 0.024;
  Matches   14;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 SGVRGDSGSLALRVARQL 20
              || ||| |:|| ||||||
Db         54 SGRRGDMGALAARVARQL 71


RESULT 10
Q67442
ID   Q67442      PRELIMINARY;      PRT;   169 AA.
AC   Q67442; Q67443;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Foot and mouth disease virus (Serotype A32), capsid protein VP1 mRNA
DE   (Fragment).
OS   Foot-and-mouth disease virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12110;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85190547; PubMed=2986125;
RA   Weddell G.N., Yansura D.G., Dowbenko D.J., Hoatlin M.E., Grubman M.J.,
RA   Moore D.M., Kleid D.G.;
RT   "Sequence variation in the gene for the immunogenic capsid protein VP1
RT   of foot-and-mouth disease virus type A.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:2618-2622(1985).
DR   EMBL; K03342; AAA42598.1; -.
DR   HSSP; Q88571; 1TME.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 2.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     169    169
SQ   SEQUENCE   169 AA;  18499 MW;  9E8AEDC94AC57F67 CRC64;

  Query Match             64.6%;  Score 64;  DB 12;  Length 169;
  Best Local Similarity   73.7%;  Pred. No. 0.033;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||| ||| | || |||:||
Db         96 GSGRRGDLGPLAARVAKQL 114


RESULT 11
Q8V442
ID   Q8V442      PRELIMINARY;      PRT;   125 AA.
AC   Q8V442;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   VP1 (Fragment).
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GAM/51/98;
RA   van Rensburg H.G., Bastos A.D., Haydon D., Joubert F., Nel L.H.;
RT   "Gene heterogeneity in the leader and 3C proteinases of foot-and-mouth
RT   disease virus.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF390862; AAL73360.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     125    125
SQ   SEQUENCE   125 AA;  13486 MW;  2D172B9F0C281A02 CRC64;

  Query Match             60.6%;  Score 60;  DB 12;  Length 125;
  Best Local Similarity   73.7%;  Pred. No. 0.1;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||  ||| ||||:||| ||
Db         53 GSTRRGDLGSLAVRVAAQL 71


RESULT 12
Q9YPU0
ID   Q9YPU0      PRELIMINARY;      PRT;   143 AA.
AC   Q9YPU0;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   VP1 protein (Fragment).
GN   VP1.
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NM/XZ/64;
RA   Liu Z., Zhao Q., Liu W., Xie Q.;
RT   "Analysis of VP1 coding nucleotide sequences of six strains of foot-
RT   and-mouth disease virus type A.";
RL   Ping Tu Hsueh Pao 14:60-67(1998).
DR   EMBL; AJ131664; CAA10474.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     143    143
SQ   SEQUENCE   143 AA;  15881 MW;  CF0CC1858CC331E0 CRC64;

  Query Match             60.6%;  Score 60;  DB 12;  Length 143;
  Best Local Similarity   72.2%;  Pred. No. 0.12;
  Matches   13;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 SGVRGDSGSLALRVARQL 20
              :| ||| |||| |||:||
Db         71 AGRRGDLGSLAARVAKQL 88


RESULT 13
Q8JUR8
ID   Q8JUR8      PRELIMINARY;      PRT;   208 AA.
AC   Q8JUR8;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
```

```
DE   1D protein (Fragment).
GN   POL.
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/IND/170/88;
RX   MEDLINE=21954626; PubMed=11958451;
RA   Tosh C., Sanyal A., Hemadri D., Venkataramanan R.;
RT   "Phylogenetic analysis of serotype A foot-and-mouth disease virus
RT   isolated in India between 1977 and 2000.";
RL   Arch. Virol. 147:493-513(2002).
DR   EMBL; AF390612; AAM63984.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  22607 MW;  6CE616F86E60CF7E CRC64;

  Query Match             60.6%;  Score 60;  DB 12;  Length 208;
  Best Local Similarity   68.4%;  Pred. No. 0.18;
  Matches   13;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              |:| ||| |:|| ||| ||
Db        140 GAGRRGDLGALAARVAAQL 158


RESULT 14
Q8JUP0
ID   Q8JUP0      PRELIMINARY;      PRT;   211 AA.
AC   Q8JUP0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   1D protein (Fragment).
GN   POL.
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/IND/302/88;
RX   MEDLINE=21954626; PubMed=11958451;
RA   Tosh C., Sanyal A., Hemadri D., Venkataramanan R.;
RT   "Phylogenetic analysis of serotype A foot-and-mouth disease virus
RT   isolated in India between 1977 and 2000.";
RL   Arch. Virol. 147:493-513(2002).
DR   EMBL; AF390641; AAM64012.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     211    211
SQ   SEQUENCE   211 AA;  23100 MW;  CFBFC94CCD9E2FE3 CRC64;

  Query Match             60.6%;  Score 60;  DB 12;  Length 211;
  Best Local Similarity   68.4%;  Pred. No. 0.18;
  Matches   13;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              |:| ||| |:|| ||| ||
Db        140 GTGRRGDLGALAARVAAQL 158


RESULT 15
Q8JUP1
ID   Q8JUP1      PRELIMINARY;      PRT;   212 AA.
AC   Q8JUP1;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   1D protein (Fragment).
GN   POL.
OS   Foot-and-mouth disease virus A.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
OX   NCBI_TaxID=12111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/IND/299/99;
RX   MEDLINE=21954626; PubMed=11958451;
RA   Tosh C., Sanyal A., Hemadri D., Venkataramanan R.;
RT   "Phylogenetic analysis of serotype A foot-and-mouth disease virus
RT   isolated in India between 1977 and 2000.";
RL   Arch. Virol. 147:493-513(2002).
DR   EMBL; AF390640; AAM64011.1; -.
DR   InterPro; IPR004080; FMDVP1coat.
DR   InterPro; IPR001676; Rhv.
DR   Pfam; PF00073; rhv; 1.
DR   PRINTS; PR01542; FMDVP1COAT.
FT   NON_TER       1      1
FT   NON_TER     212    212
SQ   SEQUENCE   212 AA;  23150 MW;  E6B690890042CD3A CRC64;

  Query Match             60.6%;  Score 60;  DB 12;  Length 212;
  Best Local Similarity   68.4%;  Pred. No. 0.18;
  Matches   13;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              |:| ||| |:|| ||| ||
Db        139 GTGRRGDLGALAARVAAQL 157


Search completed: October  9, 2003, 10:17:35
Job time : 50.3333 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:11:24 ; Search time 17.3333 Seconds
(without alignments)
48.820 Million cell updates/sec

Title: US-09-549-186B-9
Perfect score: 99
Sequence: 1 CGSGVRGDSGSLALRVARQL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	99	100.0	20	4	US-08-716-249-9	Sequence 9, Appli
2	93	93.9	20	4	US-08-716-249-8	Sequence 8, Appli
3	86	86.9	20	4	US-08-716-249-7	Sequence 7, Appli
4	86	86.9	31	1	US-08-418-716A-3	Sequence 3, Appli
5	84.5	85.4	19	4	US-08-716-249-1	Sequence 1, Appli
6	77	77.8	35	3	US-09-100-600A-1	Sequence 1, Appli
7	77	77.8	46	3	US-09-100-600A-38	Sequence 38, Appl
8	77	77.8	75	3	US-09-100-600A-28	Sequence 28, Appl
9	74	74.7	31	1	US-08-418-716A-4	Sequence 4, Appli
10	73	73.7	31	1	US-08-418-716A-5	Sequence 5, Appli
11	73	73.7	31	1	US-08-418-716A-6	Sequence 6, Appli
12	71.5	72.2	19	4	US-08-716-249-2	Sequence 2, Appli
13	71	71.7	31	1	US-08-418-716A-2	Sequence 2, Appli
14	70	70.7	31	1	US-08-418-716A-1	Sequence 1, Appli
15	70	70.7	31	1	US-08-418-716A-7	Sequence 7, Appli
16	70	70.7	46	3	US-09-100-600A-42	Sequence 42, Appl
17	70	70.7	46	3	US-09-100-600A-49	Sequence 49, Appl
18	70	70.7	46	3	US-09-100-600A-50	Sequence 50, Appl
19	69	69.7	35	3	US-09-100-600A-2	Sequence 2, Appli
20	69	69.7	35	3	US-09-100-600A-8	Sequence 8, Appli
21	69	69.7	46	3	US-09-100-600A-3	Sequence 3, Appli
22	69	69.7	65	3	US-09-100-600A-27	Sequence 27, Appl
23	69	69.7	75	3	US-09-100-600A-29	Sequence 29, Appl
24	66	66.7	46	3	US-09-100-600A-43	Sequence 43, Appl
25	66	66.7	46	3	US-09-100-600A-44	Sequence 44, Appl
26	66	66.7	46	3	US-09-100-600A-48	Sequence 48, Appl
27	64	64.6	46	3	US-09-100-600A-45	Sequence 45, Appl
28	61	61.6	35	3	US-09-100-600A-9	Sequence 9, Appli
29	61	61.6	55	3	US-09-100-600A-33	Sequence 33, Appl
30	58	58.6	46	3	US-09-100-600A-39	Sequence 39, Appl
31	57	57.6	46	3	US-09-100-600A-46	Sequence 46, Appl
32	53	53.5	45	3	US-09-100-600A-55	Sequence 55, Appl
33	53	53.5	45	3	US-09-100-600A-65	Sequence 65, Appl
34	52	52.5	45	3	US-09-100-600A-60	Sequence 60, Appl
35	52	52.5	46	3	US-09-100-600A-47	Sequence 47, Appl
36	50	50.5	170	4	US-09-732-210-1676	Sequence 1676, Ap
37	50	50.5	387	4	US-09-252-991A-17881	Sequence 17881, A
38	49	49.5	45	3	US-09-100-600A-51	Sequence 51, Appl
39	49	49.5	45	3	US-09-100-600A-52	Sequence 52, Appl
40	49	49.5	45	3	US-09-100-600A-62	Sequence 62, Appl
41	49	49.5	45	3	US-09-100-600A-64	Sequence 64, Appl
42	48	48.5	46	3	US-09-100-600A-40	Sequence 40, Appl
43	48	48.5	46	3	US-09-100-600A-41	Sequence 41, Appl
44	46	46.5	45	3	US-09-100-600A-56	Sequence 56, Appl
45	46	46.5	312	1	US-08-414-625-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-716-249-9
; Sequence 9, Application US/08716249
; Patent No. 6455244
; GENERAL INFORMATION:
; APPLICANT: Guichard, Gilles
; APPLICANT: Muller, Sylviane
; APPLICANT: Briand, Jean-Paul
; APPLICANT: Regenmortel, Marc
; TITLE OF INVENTION: Retropeptides, Antibodies Thereto, and
; TITLE OF INVENTION: Uses Thereof for Vaccination and In Vitro Diagnosis
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spencer & Frank
; STREET: 1100 New York Avenue, Suite 300E
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/716,249
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/00292
; FILING DATE: 13-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Calvetti, Frederick F.
; REGISTRATION NUMBER: 28,557
; REFERENCE/DOCKET NUMBER: GROFO 7001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)414-4000
; TELEFAX: (202)414-4040
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: YES
US-08-716-249-9

Query Match 100.0%; Score 99; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CGSGVRGDSGSLALRVARQL 20
              ||||||||||||||||||||
Db          1 CGSGVRGDSGSLALRVARQL 20


RESULT 2
US-08-716-249-8
; Sequence 8, Application US/08716249
; Patent No. 6455244
;  GENERAL INFORMATION:
;    APPLICANT:  Guichard, Gilles
;    APPLICANT:  Muller, Sylviane
;    APPLICANT:  Briand, Jean-Paul
;    APPLICANT:  Regenmortel, Marc
;    TITLE OF INVENTION:  Retropeptides, Antibodies Thereto, and
;    TITLE OF INVENTION:  Uses Thereof for Vaccination and In Vitro Diagnosis
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spencer & Frank
;      STREET:  1100 New York Avenue, Suite 300E
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/716,249
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/00292
;      FILING DATE:  13-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Calvetti, Frederick F.
;      REGISTRATION NUMBER:  28,557
;      REFERENCE/DOCKET NUMBER:  GROFO 7001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)414-4000
;      TELEFAX:  (202)414-4040
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  YES
US-08-716-249-8

  Query Match             93.9%;  Score 93;  DB 4;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 8.7e-08;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDSGSLALRVARQL 20
              ||||||||| ||||||||||
Db          1 CGSGVRGDFGSLALRVARQL 20


RESULT 3
US-08-716-249-7
; Sequence 7, Application US/08716249
; Patent No. 6455244
;  GENERAL INFORMATION:
;    APPLICANT:  Guichard, Gilles
;    APPLICANT:  Muller, Sylviane
;    APPLICANT:  Briand, Jean-Paul
;    APPLICANT:  Regenmortel, Marc
;    TITLE OF INVENTION:  Retropeptides, Antibodies Thereto, and
;    TITLE OF INVENTION:  Uses Thereof for Vaccination and In Vitro Diagnosis
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spencer & Frank
;      STREET:  1100 New York Avenue, Suite 300E
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/716,249
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/00292
;      FILING DATE:  13-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Calvetti, Frederick F.
;      REGISTRATION NUMBER:  28,557
;      REFERENCE/DOCKET NUMBER:  GROFO 7001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)414-4000
;      TELEFAX:  (202)414-4040
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-716-249-7

  Query Match             86.9%;  Score 86;  DB 4;  Length 20;
  Best Local Similarity   90.0%;  Pred. No. 9.8e-07;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDSGSLALRVARQL 20
              |||||||| |||| ||||||
Db          1 CGSGVRGDFGSLAPRVARQL 20


RESULT 4
US-08-418-716A-3
; Sequence 3, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;    APPLICANT:  Mason, Peter W.
;    APPLICANT:  Baxt, Barry
;    APPLICANT:  Reider, Elizabeth
;    APPLICANT:  Berinstein, Analia
;    APPLICANT:  Kang, Angray S
;    TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;    TITLE OF INVENTION:  Viruses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janelle S. Graeter
;      STREET:  Rm. 411, Bldg. 005, BARC-W
;      CITY:  Beltsville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20705
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/418,716A
;      FILING DATE:
```

```
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Graeter, Janelle S.
;      REGISTRATION NUMBER:  35,024
;      REFERENCE/DOCKET NUMBER:  0137.94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 504-6629
;      TELEFAX:  (301) 504-5060
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:
;      ORGANISM:  Foot and mouth disease virus
US-08-418-716A-3

  Query Match             86.9%;  Score 86;  DB 1;  Length 31;
  Best Local Similarity   94.7%;  Pred. No. 1.6e-06;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||||| ||||||||||
Db         11 GSGVRGDKGSLALRVARQL 29


RESULT 5
US-08-716-249-1
; Sequence 1, Application US/08716249
; Patent No. 6455244
;  GENERAL INFORMATION:
;    APPLICANT:  Guichard, Gilles
;    APPLICANT:  Muller, Sylviane
;    APPLICANT:  Briand, Jean-Paul
;    APPLICANT:  Regenmortel, Marc
;    TITLE OF INVENTION:  Retropeptides, Antibodies Thereto, and
;    TITLE OF INVENTION:  Uses Thereof for Vaccination and In Vitro Diagnosis
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spencer & Frank
;      STREET:  1100 New York Avenue, Suite 300E
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/716,249
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/00292
;      FILING DATE:  13-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Calvetti, Frederick F.
;      REGISTRATION NUMBER:  28,557
;      REFERENCE/DOCKET NUMBER:  GROFO 7001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)414-4000
;      TELEFAX:  (202)414-4040
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-716-249-1

  Query Match             85.4%;  Score 84.5;  DB 4;  Length 19;
  Best Local Similarity   95.0%;  Pred. No. 1.6e-06;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 CGSGVRGDSGSLALRVARQL 20
              . |||||||||| ||||||||
Db          1 CGSGVRGDSGS-ALRVARQL 19


RESULT 6
US-09-100-600A-1
; Sequence 1, Application US/09100600A
; Patent No. 6107021
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Shen, Ming
;    TITLE OF INVENTION:  SYNTHETIC PEPTIDE VACCINES FOR
;    TITLE OF INVENTION:  FOOT-AND-MOUTH DISEASE
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,600A
;      FILING DATE:  20-Jun-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria C.H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4156
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  35 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-600A-1

  Query Match             77.8%;  Score 77;  DB 3;  Length 35;
  Best Local Similarity   89.5%;  Pred. No. 4e-05;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||||| |||| ||||||
Db          7 GSGVRGDFGSLAPRVARQL 25


RESULT 7
US-09-100-600A-38
; Sequence 38, Application US/09100600A
; Patent No. 6107021
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Shen, Ming
;    TITLE OF INVENTION:  SYNTHETIC PEPTIDE VACCINES FOR
;    TITLE OF INVENTION:  FOOT-AND-MOUTH DISEASE
;    NUMBER OF SEQUENCES:  91
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,600A
;      FILING DATE:  20-Jun-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria C.H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4156
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-600A-38

  Query Match             77.8%;  Score 77;  DB 3;  Length 46;
  Best Local Similarity   89.5%;  Pred. No. 5.3e-05;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||||| |||| ||||||
Db         16 GSGVRGDFGSLAPRVARQL 34


RESULT 8
US-09-100-600A-28
; Sequence 28, Application US/09100600A
; Patent No. 6107021
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Shen, Ming
;    TITLE OF INVENTION:  SYNTHETIC PEPTIDE VACCINES FOR
;    TITLE OF INVENTION:  FOOT-AND-MOUTH DISEASE
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,600A
;      FILING DATE:  20-Jun-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria C.H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4156
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  75 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-600A-28

  Query Match             77.8%;  Score 77;  DB 3;  Length 75;
  Best Local Similarity   89.5%;  Pred. No. 8.9e-05;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||||| |||| ||||||
Db         47 GSGVRGDFGSLAPRVARQL 65


RESULT 9
US-08-418-716A-4
; Sequence 4, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;    APPLICANT:  Mason, Peter W.
;    APPLICANT:  Baxt, Barry
;    APPLICANT:  Reider, Elizabeth
;    APPLICANT:  Berinstein, Analia
;    APPLICANT:  Kang, Angray S
;    TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;    TITLE OF INVENTION:  Viruses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janelle S. Graeter
;      STREET:  Rm. 411, Bldg. 005, BARC-W
;      CITY:  Beltsville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20705
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/418,716A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Graeter, Janelle S.
;      REGISTRATION NUMBER:  35,024
;      REFERENCE/DOCKET NUMBER:  0137.94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 504-6629
;      TELEFAX:  (301) 504-5060
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:
;      ORGANISM:  Foot and mouth disease virus
US-08-418-716A-4

  Query Match             74.7%;  Score 74;  DB 1;  Length 31;
  Best Local Similarity   84.2%;  Pred. No. 0.0001;
  Matches   16;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||:|| |||| ||||||
Db         11 GSGVKGDFGSLAPRVARQL 29
```

```
RESULT 10
US-08-418-716A-5
; Sequence 5, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;    APPLICANT:  Mason, Peter W.
;    APPLICANT:  Baxt, Barry
;    APPLICANT:  Reider, Elizabeth
;    APPLICANT:  Berinstein, Analia
;    APPLICANT:  Kang, Angray S
;    TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;    TITLE OF INVENTION:  Viruses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janelle S. Graeter
;      STREET:  Rm. 411, Bldg. 005, BARC-W
;      CITY:  Beltsville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20705
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/418,716A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Graeter, Janelle S.
;      REGISTRATION NUMBER:  35,024
;      REFERENCE/DOCKET NUMBER:  0137.94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 504-6629
;      TELEFAX:  (301) 504-5060
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:
;      ORGANISM:  Foot and mouth disease virus
US-08-418-716A-5

  Query Match             73.7%;  Score 73;  DB 1;  Length 31;
  Best Local Similarity   84.2%;  Pred. No. 0.00014;
  Matches   16;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||||: |||| ||||||
Db         11 GSGVRGEFGSLAPRVARQL 29


RESULT 11
US-08-418-716A-6
; Sequence 6, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;    APPLICANT:  Mason, Peter W.
;    APPLICANT:  Baxt, Barry
;    APPLICANT:  Reider, Elizabeth
;    APPLICANT:  Berinstein, Analia
;    APPLICANT:  Kang, Angray S
;    TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;    TITLE OF INVENTION:  Viruses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janelle S. Graeter
;      STREET:  Rm. 411, Bldg. 005, BARC-W
;      CITY:  Beltsville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20705
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/418,716A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Graeter, Janelle S.
;      REGISTRATION NUMBER:  35,024
;      REFERENCE/DOCKET NUMBER:  0137.94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 504-6629
;      TELEFAX:  (301) 504-5060
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:
;      ORGANISM:  Foot and mouth disease virus
US-08-418-716A-6

  Query Match             73.7%;  Score 73;  DB 1;  Length 31;
  Best Local Similarity   84.2%;  Pred. No. 0.00014;
  Matches   16;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||||: |||| ||||||
Db         11 GSGVRGEFGSLAPRVARQL 29


RESULT 12
US-08-716-249-2
; Sequence 2, Application US/08716249
; Patent No. 6455244
;  GENERAL INFORMATION:
;    APPLICANT:  Guichard, Gilles
;    APPLICANT:  Muller, Sylviane
;    APPLICANT:  Briand, Jean-Paul
;    APPLICANT:  Regenmortel, Marc
;    TITLE OF INVENTION:  Retropeptides, Antibodies Thereto, and
;    TITLE OF INVENTION:  Uses Thereof for Vaccination and In Vitro Diagnosis
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spencer & Frank
;      STREET:  1100 New York Avenue, Suite 300E
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/716,249
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  WO PCT/FR95/00292
;       FILING DATE:  13-MAR-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Calvetti, Frederick F.
;       REGISTRATION NUMBER:  28,557
;       REFERENCE/DOCKET NUMBER:  GROFO 7001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202)414-4000
;       TELEFAX:  (202)414-4040
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  19 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
US-08-716-249-2

  Query Match             72.2%;  Score 71.5;  DB 4;  Length 19;
  Best Local Similarity   85.0%;  Pred. No. 0.00014;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 CGSGVRGDSGSLALRVARQL 20
              |||||||| || | ||||||
Db          1 CGSGVRGDFGS-APRVARQL 19


RESULT 13
US-08-418-716A-2
; Sequence 2, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;     APPLICANT:  Mason, Peter W.
;     APPLICANT:  Baxt, Barry
;     APPLICANT:  Reider, Elizabeth
;     APPLICANT:  Berinstein, Analia
;     APPLICANT:  Kang, Angray S
;     TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;     TITLE OF INVENTION:  Viruses
;     NUMBER OF SEQUENCES:  8
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Janelle S. Graeter
;       STREET:  Rm. 411, Bldg. 005, BARC-W
;       CITY:  Beltsville
;       STATE:  Maryland
;       COUNTRY:  USA
;       ZIP:  20705
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/418,716A
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Graeter, Janelle S.
;       REGISTRATION NUMBER:  35,024
;       REFERENCE/DOCKET NUMBER:  0137.94
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 504-6629
;       TELEFAX:  (301) 504-5060
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  31 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     HYPOTHETICAL:  NO
;     FRAGMENT TYPE:  internal
;     ORIGINAL SOURCE:
;       ORGANISM:  Foot and mouth disease virus
US-08-418-716A-2

  Query Match             71.7%;  Score 71;  DB 1;  Length 31;
  Best Local Similarity   84.2%;  Pred. No. 0.00028;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||| ||| |||| ||||||
Db         11 GSGPRGDFGSLAPRVARQL 29


RESULT 14
US-08-418-716A-1
; Sequence 1, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;     APPLICANT:  Mason, Peter W.
;     APPLICANT:  Baxt, Barry
;     APPLICANT:  Reider, Elizabeth
;     APPLICANT:  Berinstein, Analia
;     APPLICANT:  Kang, Angray S
;     TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;     TITLE OF INVENTION:  Viruses
;     NUMBER OF SEQUENCES:  8
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Janelle S. Graeter
;       STREET:  Rm. 411, Bldg. 005, BARC-W
;       CITY:  Beltsville
;       STATE:  Maryland
;       COUNTRY:  USA
;       ZIP:  20705
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/418,716A
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Graeter, Janelle S.
;       REGISTRATION NUMBER:  35,024
;       REFERENCE/DOCKET NUMBER:  0137.94
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 504-6629
;       TELEFAX:  (301) 504-5060
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  31 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     HYPOTHETICAL:  NO
;     FRAGMENT TYPE:  internal
;     ORIGINAL SOURCE:
;       ORGANISM:  Foot and mouth disease virus
US-08-418-716A-1

  Query Match             70.7%;  Score 70;  DB 1;  Length 31;
  Best Local Similarity   84.2%;  Pred. No. 0.0004;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||| ||| |||| ||||||
Db         11 GSGDRGDFGSLAPRVARQL 29


RESULT 15
US-08-418-716A-7
```

```
; Sequence 7, Application US/08418716A
; Patent No. 5612040
;  GENERAL INFORMATION:
;    APPLICANT:  Mason, Peter W.
;    APPLICANT:  Baxt, Barry
;    APPLICANT:  Reider, Elizabeth
;    APPLICANT:  Berinstein, Analia
;    APPLICANT:  Kang, Angray S
;    TITLE OF INVENTION:  No. 5612040-Infectious Foot-and-Mouth Disease
;    TITLE OF INVENTION:  Viruses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Janelle S. Graeter
;      STREET:  Rm. 411, Bldg. 005, BARC-W
;      CITY:  Beltsville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20705
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/418,716A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Graeter, Janelle S.
;      REGISTRATION NUMBER:  35,024
;      REFERENCE/DOCKET NUMBER:  0137.94
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 504-6629
;      TELEFAX:  (301) 504-5060
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:
;      ORGANISM:  Foot and mouth disease virus
US-08-418-716A-7

  Query Match            70.7%;  Score 70;  DB 1;  Length 31;
  Best Local Similarity  78.9%;  Pred. No. 0.0004;
  Matches   15;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              ||||:|: |||| ||||||
Db         11 GSGVKGEFGSLAPRVARQL 29


Search completed: October  9, 2003, 10:19:38
Job time : 17.3333 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:14:31 ; Search time 21.3333 Seconds
(without alignments)
151.058 Million cell updates/sec

Title: US-09-549-186B-9
Perfect score: 99
Sequence: 1 CGSGVRGDSGSLALRVARQL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	50	50.5	173	15	US-10-156-761-12492	Sequence 12492, A
2	46	46.5	312	15	US-10-207-655-186	Sequence 186, App
3	46	46.5	559	9	US-09-988-200-2	Sequence 2, Appli
4	45	45.5	1208	15	US-10-156-761-13251	Sequence 13251, A
5	44	44.4	74	15	US-10-106-698-7642	Sequence 7642, Ap
6	43	43.4	294	15	US-10-102-806-595	Sequence 595, App
7	42	42.4	508	11	US-09-895-298-92	Sequence 92, Appl
8	42	42.4	1014	9	US-09-912-020-266	Sequence 266, App
9	41	41.4	87	10	US-09-764-877-1521	Sequence 1521, Ap
10	41	41.4	261	15	US-10-156-761-9294	Sequence 9294, Ap
11	41	41.4	449	10	US-09-919-497-89	Sequence 89, Appl
12	41	41.4	458	9	US-09-925-301-1282	Sequence 1282, Ap
13	41	41.4	459	10	US-09-925-300-1440	Sequence 1440, Ap
14	41	41.4	532	11	US-09-941-947A-38	Sequence 38, Appl
15	41	41.4	532	15	US-10-209-372-2	Sequence 2, Appli
16	41	41.4	722	14	US-10-128-870-23	Sequence 23, Appl
17	41	41.4	722	15	US-10-131-685-23	Sequence 23, Appl
18	41	41.4	3500	15	US-10-153-219-2	Sequence 2, Appli
19	41	41.4	3537	15	US-10-153-219-15	Sequence 15, Appl
20	40.5	40.9	378	12	US-10-166-225A-49	Sequence 49, Appl
21	40	40.4	182	10	US-09-738-626-4147	Sequence 4147, Ap
22	40	40.4	216	9	US-09-908-180-4	Sequence 4, Appli
23	40	40.4	274	10	US-09-738-626-6076	Sequence 6076, Ap
24	40	40.4	283	15	US-10-156-761-9949	Sequence 9949, Ap
25	40	40.4	316	15	US-10-106-698-5711	Sequence 5711, Ap
26	40	40.4	337	9	US-09-815-242-11926	Sequence 11926, A
27	40	40.4	350	11	US-09-847-208-53	Sequence 53, Appl
28	40	40.4	390	9	US-09-925-297-531	Sequence 531, App
29	40	40.4	476	9	US-09-291-299A-3	Sequence 3, Appli
30	40	40.4	627	15	US-10-222-100-3	Sequence 3, Appli
31	40	40.4	784	9	US-09-905-983-52	Sequence 52, Appl
32	40	40.4	784	12	US-10-021-660-99	Sequence 99, Appl
33	40	40.4	865	11	US-09-842-758-20	Sequence 20, Appl
34	40	40.4	934	11	US-09-842-758-18	Sequence 18, Appl
35	40	40.4	1783	12	US-10-276-934-12	Sequence 12, Appl
36	40	40.4	1800	12	US-10-276-934-10	Sequence 10, Appl
37	40	40.4	1826	12	US-10-276-934-9	Sequence 9, Appli
38	40	40.4	2008	12	US-10-276-934-11	Sequence 11, Appl
39	40	40.4	2306	12	US-10-276-934-14	Sequence 14, Appl
40	40	40.4	2352	12	US-10-276-934-13	Sequence 13, Appl
41	39.5	39.9	446	10	US-09-738-626-6262	Sequence 6262, Ap
42	39.5	39.9	531	15	US-10-156-761-13662	Sequence 13662, A
43	39	39.4	89	9	US-09-925-301-1058	Sequence 1058, Ap
44	39	39.4	89	15	US-10-106-698-5417	Sequence 5417, Ap
45	39	39.4	124	16	US-10-080-170-376	Sequence 376, App

ALIGNMENTS

RESULT 1
US-10-156-761-12492
; Sequence 12492, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12492
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12492

Query Match 50.5%; Score 50; DB 15; Length 173;
Best Local Similarity 57.9%; Pred. No. 5.3;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```
Qy          2 GSGVRGDSGSLALRVARQL 20
              | || | :|:| | ||| |
Db        111 GGGVSGQAGALRLGVARAL 129
```

RESULT 2
US-10-207-655-186

```
; Sequence 186, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
;  APPLICANT: Ledbetter, Jeffrey A.
;  APPLICANT:  Hayden-Ledbetter, Martha S.
;  TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
;  FILE REFERENCE: 390069.401C1
;  CURRENT APPLICATION NUMBER: US/10/207,655
;  CURRENT FILING DATE:  2002-07-25
;  NUMBER OF SEQ ID NOS: 426
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 186
;   LENGTH: 312
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-207-655-186

  Query Match             46.5%;  Score 46;  DB 15;  Length 312;
  Best Local Similarity   61.5%;  Pred. No. 40;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDSGSLA 13
              |||| ||  | :|
Db        164 CGSGARGGDGEVA 176


RESULT 3
US-09-988-200-2
; Sequence 2, Application US/09988200
; Patent No. US20020094553A1
;   GENERAL INFORMATION:
;        APPLICANT: BARBEYRON, Tristan
;                   POTIN, Philippe
;                   RICHARD, Christophe
;                   HENRISSAT, Bernard
;                   YVIN, Jean-Claude
;                   KLOAREG, Bernard
;        TITLE OF INVENTION: Glycolyse hydrolase genes and their
;                            use for producing enzymes for the biodegradation of
;                            carrageenans
;        NUMBER OF SEQUENCES: 8
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DENNISON, MESEROLE, SCHEINER & SCHULTZ
;             STREET: 612 Crystal Square 4, 1745 Jefferson Davis
;                     Highway
;             CITY: ARLINGTON
;             STATE: VIRGINIA
;             COUNTRY: U.S.A.
;             ZIP: 22202
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/988,200
;             FILING DATE: 19-No. US20020094553A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/269,731
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: FR 96 12204
;             FILING DATE: 07-OCT-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: IRA SCHULTZ
;             REGISTRATION NUMBER: <Unknown>
;             REFERENCE/DOCKET NUMBER: <Unknown>
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (703) 412-1155
;             TELEFAX: (703) 412-1161
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 559 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-988-200-2

  Query Match             46.5%;  Score 46;  DB 9;  Length 559;
  Best Local Similarity   64.3%;  Pred. No. 75;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDSGSLAL 14
              ||| || ||| : |
Db        298 CGSAVRSDSGFVEL 311


RESULT 4
US-10-156-761-13251
; Sequence 13251, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13251
;   LENGTH: 1208
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-13251

  Query Match             45.5%;  Score 45;  DB 15;  Length 1208;
  Best Local Similarity   81.8%;  Pred. No. 2.5e+02;
  Matches    9;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 GDSGSLALRVA 17
              ||||:||||:|
Db        623 GDSGTLALRLA 633


RESULT 5
US-10-106-698-7642
; Sequence 7642, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept
;  FILE REFERENCE: PA005P1
;  CURRENT APPLICATION NUMBER: US/10/106,698
;  CURRENT FILING DATE:  2002-03-27
;  PRIOR APPLICATION NUMBER: PCT/US00/26524
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: US 60/157,137
;  PRIOR FILING DATE: 1999-09-29
;  PRIOR APPLICATION NUMBER: US 60/163,280
;  PRIOR FILING DATE: 1999-11-03
;  NUMBER OF SEQ ID NOS: 8564
;  SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 7642
;   LENGTH: 74
;   TYPE: PRT
;   ORGANISM: Homo sapiens
```

```
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (49)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (54)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-7642

  Query Match             44.4%;  Score 44;  DB 15;  Length 74;
  Best Local Similarity   57.9%;  Pred. No. 17;
  Matches   11;  Conservative    1;  Mismatches    5;  Indels    2;  Gaps    1;

Qy          1 CGSG--VRGDSGSLALRVA 17
              |||   : |  ||||| ||
Db         26 CGSSRDLHGSQGSLALSVA 44


RESULT 6
US-10-102-806-595
; Sequence 595, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA103P1C1
;  CURRENT APPLICATION NUMBER: US/10/102,806
;  CURRENT FILING DATE:  2002-03-22
;  PRIOR APPLICATION NUMBER: 09/925,298
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05881
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 846
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 595
;   LENGTH: 294
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (269)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (278)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-595

  Query Match             43.4%;  Score 43;  DB 15;  Length 294;
  Best Local Similarity   53.3%;  Pred. No. 1.1e+02;
  Matches    8;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 SGVRGDSGSLALRVA 17
              ||  || | | :|:|
Db        159 SGTSGDHGELVVRIA 173


RESULT 7
US-09-895-298-92
; Sequence 92, Application US/09895298
; Publication No. US20030078405A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 47 Human Secreted Proteins
;  FILE REFERENCE: PZ035P1
;  CURRENT APPLICATION NUMBER: US/09/895,298
;  CURRENT FILING DATE:  2001-07-02
;  PRIOR APPLICATION NUMBER: 09/591,16
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: PCT/US99/29950
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: 60/113,006
;  PRIOR FILING DATE: 1998-12-18
;  PRIOR APPLICATION NUMBER: 60/112,809
;  PRIOR FILING DATE: 1998-12-17
;  NUMBER OF SEQ ID NOS: 231
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 92
;   LENGTH: 508
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-895-298-92

  Query Match             42.4%;  Score 42;  DB 11;  Length 508;
  Best Local Similarity   56.2%;  Pred. No. 2.8e+02;
  Matches    9;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDSGSLALRV 16
              ||||   | |  || |
Db        215 CGSGEENDPGEQALPV 230


RESULT 8
US-09-912-020-266
; Sequence 266, Application US/09912020
; Patent No. US20020045592A1
; GENERAL INFORMATION:
;  APPLICANT: Zyskind, Judith
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Trawick, John
;  APPLICANT:  Forsyth, R. Allyn
;  APPLICANT:  Froelich, Jamie M.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
;  TITLE OF INVENTION:  ESCHERICHIA COLI
;  FILE REFERENCE: ELITRA.001DV1
;  CURRENT APPLICATION NUMBER: US/09/912,020
;  CURRENT FILING DATE:  2001-07-23
;  PRIOR APPLICATION NUMBER: 09/492,709
;  PRIOR FILING DATE: 2000-01-27
;  PRIOR APPLICATION NUMBER: 60/117,405
;  PRIOR FILING DATE: 1999-01-27
;  NUMBER OF SEQ ID NOS: 485
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 266
;   LENGTH: 1014
;   TYPE: PRT
;   ORGANISM: E. Coli
US-09-912-020-266

  Query Match             42.4%;  Score 42;  DB 9;  Length 1014;
  Best Local Similarity   40.0%;  Pred. No. 5.8e+02;
  Matches    8;  Conservative    3;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDSGSLALRVARQL 20
              | ||   ::| |  : || |
Db        165 CASGASNETGMLTQKFARSL 184


RESULT 9
US-09-764-877-1521
; Sequence 1521, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC005
;  CURRENT APPLICATION NUMBER: US/09/764,877
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 4031
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1521
```

```
;   LENGTH: 87
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-764-877-1521

  Query Match             41.4%;  Score 41;  DB 10;  Length 87;
  Best Local Similarity   88.9%;  Pred. No. 58;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSG 10
              ||| |||||
Db         22 GSGKRGDSG 30


RESULT 10
US-10-156-761-9294
; Sequence 9294, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9294
;   LENGTH: 261
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-9294

  Query Match             41.4%;  Score 41;  DB 15;  Length 261;
  Best Local Similarity   52.6%;  Pred. No. 1.9e+02;
  Matches   10;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVARQL 20
              |  |  | |: | |:||||
Db         97 GQSVLLDIGTTAYRLARQL 115


RESULT 11
US-09-919-497-89
; Sequence 89, Application US/09919497
; Patent No. US20020106662A1
; GENERAL INFORMATION:
;  APPLICANT: Mutter, George L.
;  TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
;  FILE REFERENCE: B0801/7225
;  CURRENT APPLICATION NUMBER: US/09/919,497
;  CURRENT FILING DATE:  2001-07-31
;  PRIOR APPLICATION NUMBER: US 60/221,735
;  PRIOR FILING DATE: 2000-07-31
;  NUMBER OF SEQ ID NOS: 100
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 89
;   LENGTH: 449
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-919-497-89

  Query Match             41.4%;  Score 41;  DB 10;  Length 449;
  Best Local Similarity   53.8%;  Pred. No. 3.4e+02;
  Matches    7;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDSGSLA 13
              ||  |:|:|| :|
Db         37 CGGDVKGESGYVA 49


RESULT 12
US-09-925-301-1282
; Sequence 1282, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1282
;   LENGTH: 458
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (249)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1282

  Query Match             41.4%;  Score 41;  DB 9;  Length 458;
  Best Local Similarity   53.8%;  Pred. No. 3.5e+02;
  Matches    7;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CGSGVRGDSGSLA 13
              ||  |:|:|| :|
Db         46 CGGDVKGESGYVA 58


RESULT 13
US-09-925-300-1440
; Sequence 1440, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
;  APPLICANT: Craig Rosen,
;  APPLICANT:  Steve Ruben
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA101
;  CURRENT APPLICATION NUMBER: US/09/925,300
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05988
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1890
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1440
;   LENGTH: 459
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-300-1440

  Query Match             41.4%;  Score 41;  DB 10;  Length 459;
  Best Local Similarity   62.5%;  Pred. No. 3.5e+02;
  Matches   10;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          5 VRGDSGSLALRVARQL 20
              ||| ||: | | || |
Db         61 VRGGSGASARRSARSL 76
```

```
RESULT 14
US-09-941-947A-38
; Sequence 38, Application US/09941947A
; Publication No. US20030003528A1
; GENERAL INFORMATION:
;  APPLICANT: Brzostowicz, Patricia C.
;  APPLICANT:  Cheng, Qiong
;  APPLICANT:  DiCosimo, Deana J.
;  APPLICANT:  Koffas, Mattheos
;  APPLICANT:  Miller, Edward S. Jr.
;  APPLICANT:  Odom, J. Martin
;  APPLICANT:  Picataggio, Steve
;  APPLICANT:  Rouviere, Pierre E.
;  TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
;  FILE REFERENCE: CL1903 US NA
;  CURRENT APPLICATION NUMBER: US/09/941,947A
;  CURRENT FILING DATE:  2001-09-01
;  PRIOR APPLICATION NUMBER: 60/229,907
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/229,858
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 38
;   LENGTH: 532
;   TYPE: PRT
;   ORGANISM: Rhodococcus erythropolis AN12
US-09-941-947A-38

  Query Match             41.4%;  Score 41;  DB 11;  Length 532;
  Best Local Similarity   56.2%;  Pred. No. 4.1e+02;
  Matches    9;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVA 17
              | || | ||  | |:|
Db        498 GGGVSGASGRSAARIA 513


RESULT 15
US-10-209-372-2
; Sequence 2, Application US/10209372
; Publication No. US20030100045A1
; GENERAL INFORMATION:
;  APPLICANT: E. I . du Pont de Nemours, Inc.
;  APPLICANT:  Cheng, Qiong
;  APPLICANT:  Tao, Luan
;  TITLE OF INVENTION: CAROTENOID KETOLASE GENE
;  FILE REFERENCE: CL-1849 US NA
;  CURRENT APPLICATION NUMBER: US/10/209,372
;  CURRENT FILING DATE:  2002-07-30
;  PRIOR APPLICATION NUMBER:
;  PRIOR FILING DATE:
;  NUMBER OF SEQ ID NOS: 47
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
;   LENGTH: 532
;   TYPE: PRT
;   ORGANISM: Rhodococcus erythropolis AN12
US-10-209-372-2

  Query Match             41.4%;  Score 41;  DB 15;  Length 532;
  Best Local Similarity   56.2%;  Pred. No. 4.1e+02;
  Matches    9;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 GSGVRGDSGSLALRVA 17
              | ||.| ||  | |:|
Db        498 GGGVSGASGRSAARIA 513


Search completed: October  9, 2003, 10:20:49
Job time : 22.3333 secs
```